THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Lois Cronky Mullin

TITLE OF THESIS: An empirical test of the bisymmetry axiom using the method of constant stimuli.

DEGREE FOR WHICH THESIS WAS PRESENTED: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1979

THE UNIVERSITY OF ALBERTA

AN EMPIRICAL TEST OF THE BISYMMETRY AXIOM USING THE METHOD OF CONSTANT STIMULI

by

© Lois Cronky Mullin

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1979

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled An empirical test of the bisymmetry axiom using the method of constant stimuli submitted by Lois Cronky Mullin in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

In a bisection task, the subject is presented with two standard stimuli and he is required to select or adjust a third, variable stimulus that appears to be equally distant from each of the standards. Given that a subject is capable of equating subjective intervals, the data from such an experiment should provide a bsis for constructing a scale of sensation with equal interval properties. A number of researchers have been concerned with the properties of a bisymmetric structure from the axiomatic approach of measurement theory. The axioms of a bisymmetric structure are existence, monotonicity, continuity and bisymmetry. The bisymmetry axiom is the target of empirical verification since the other three are technical axioms assumed to be true given a finely graded stimulus continuum. However, most researchers have been concerned with investigating the bisection task to test non-parametric scalability, a condition which requires the additional axioms of reflexivity and commutativity to be satisfied.

To date there has been only one direct test of the bisymmetric structure without the added axioms. Cross (1968) found support for the bisymmetry axiom, but there were some difficulties with his procedure. He used the method of adjustment but did not adequately control for possible biases in the data due to the order in which the stimuli were presented and the initial intensity of the adjustable stimulus.

The present research used a different methodology, the method of constant stimuli, to directly test the validity of the bisymmetry axiom and controlled for sources of bias in the data arising from the order of presentation of the stimuli. Subjects judged whether a variable stimulus was subjectively above or below the midpoint of the two endpoint standards by pressing one of two buttons. The choice response and the latency of that response were obtained so that the verification of the bisymmetry axiom would be based on converging evidence. The endpoint standards were presented sequentially in both ascending and descending orders to control for what Stevens (1957) has called hysteresis effects.

The midpoints obtained from an analysis of the frequency data for ascending and descending trials were different, indicating that hysteresis effects were present. However, the bisection points needed for a test of the bisymmetry axiom were equal within experimental error suggesting that prior attempts to test the bisection model could be attributable to a violation of reflexivity or commutativity, or other biases of methodology. Concurrent reaction time data followed the hypothesized inverted U-shaped function, but displayed complex interactions that appear to be a function of the absolute intensity of the stimuli and the order of presentation of the endpoint standards and variable stimuli in any one trial.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

The assumptions that sensations can be measured and that their measurement scales relate to sensory functioning in some manner more fundamental than merely their heuristic use has been controversial since long before the time of Fechner. They are still. Fechner thought he had answered the questions when he proposed that a scale of sensation could be obtained by cumulating just noticeable differences between stimuli along the stimulus continuum. He advocated that the relation between the physical intensity of the stimuli and their subjective correlates was a logarithmic function.

But, almost 100 years later, S. S. Stevens argued that scales constructed from Fechnerian procedures are not valid because just noticeable differences are not subjectively equal to each other over the entire stimulus continuum. He objected to turning what he called confusions among stimuli into units of measurement, and maintained that direct estimations of subjective magnitude provide an orderly relation with stimulus intensity, a relation described by a power function.

The issue of sensory measurement remains as a fundamental question in psychology. More recently, theories and models of measurement have received renewed interest. In particular, the volume "Foundations of Measurement" by Krantz, Luce, Suppes and Tversky (1971) has brought these issues into the view of psychologists with some impact.

It presented the axioms and theorems for a great variety of measurement structures, pointed to the crucial axioms requiring empirical verification, and reviewed the relevant empirical accomplishments. Unfortunately, empirical verification of the testable axioms has received less attention then what would seem warranted by their importance.

Krantz, et al. (1971) pointed out some of the difficulties in empirically verifying an axiom system. The most obvious is errors of measurement that can lead to the failure to verify the minimum requirement of most measurement models, a weak ordering of the data. Another difficulty is introduced when generalizations are made from a small, finite set of stimulus values used in the empirical test to the infinite set encompassed in the statement of the model. The use of finite sets can violate the axioms that are assumed to be true without empirical test such as continuity.

One axiomatic system that has received empirical investigation is the bisection system, or what Krantz, et al. referred to more generally as a bisymmetric structure. In Stevens' terms, the bisection paradigm is a direct procedure in which it is assumed that subjects are able to equate subjective sensory intervals. Because previous attempts to empirically validate the system have lacked rigor in many particulars, the question remains open. Before turning to a discussion of the empirical work, however, the axiomatic foundations of bisection are presented.

Axioms and Theorems of the Bisection Model

Pfanzagl (1968) and, in a mathematically more sophisticated form, Krantz, et al. (1971) presented the axioms and proved the theorems of a bisection model. The following presentation is from those authors and from Fagot and Stewart's (1970) treatment of Pfanzagl's development of the model.

Let P denote the set of stimuli, weakly ordered by $\geq$ on the physical scale.

1. Existence Axiom. For all a, b in P, there exists a unique element (a o b) in P, which is interpreted as the bisection point of a and b.

2. Bisymmetry Axiom. For all a, b, c and d in P, (a o b) o (c o d) = (a o c) o (b o d).

3. Monotonicity Axiom. If a $\gtrless$ b, then for all c in P, (a o c) $\gtrless$ (b o c).

4. Continuity Axiom. The element (a o b) is a continuous function of both a and b.

Pfanzagl showed that if Axioms 1 through 4 are satisfied, then there exists a function, F, and real numbers p, q, and r, such that:

$$F(a \circ b) = pF(a) + qF(b) + r. \qquad (1)$$

He further showed that this representation is unique up to a positive linear transformation, i.e., the scale is

an interval scale.

Other axioms may be added to the model. For example, the following reflexivity axiom provides for increased specification of Equation 1.

5. Reflexivity Axiom. For all a in P, (a o a) = a.

Pfanzagl showed that p + q = 1 and r = 0 if and only if bisection is reflexive. Thus, Equation 1 reduces to:

$$F(a \circ b) = pF(a) + (1 - p)F(b). \qquad (2)$$

Given a stimulus continuum that is finely graded, the existence, reflexivity, monotonicity and continuity axioms are usually assumed to hold and are not the focus of empirical verification. The existence axiom defines the bisection operation "o" and the continuity axiom has no empirically testable consequences. Monotonicity is usually not violated except when c is very large in relation to a and b and there is a failure in sensitiviey or discrimination.

The crucial axiom for the empirical test of the model is the bisymmetry axiom. This axiom is presented schematically in Figure 1. Given any four standard stimuli a, b, c, and d, not necessarily equally spaced on P, the intervals from adjacent standards, (a,b) and (c,d) may be bisected to obtain bisection points (a o b) and (c o d). The interval defined by these bisection points may then be bisected to obtain (a o b) o (c o d). This sequence of bisections is

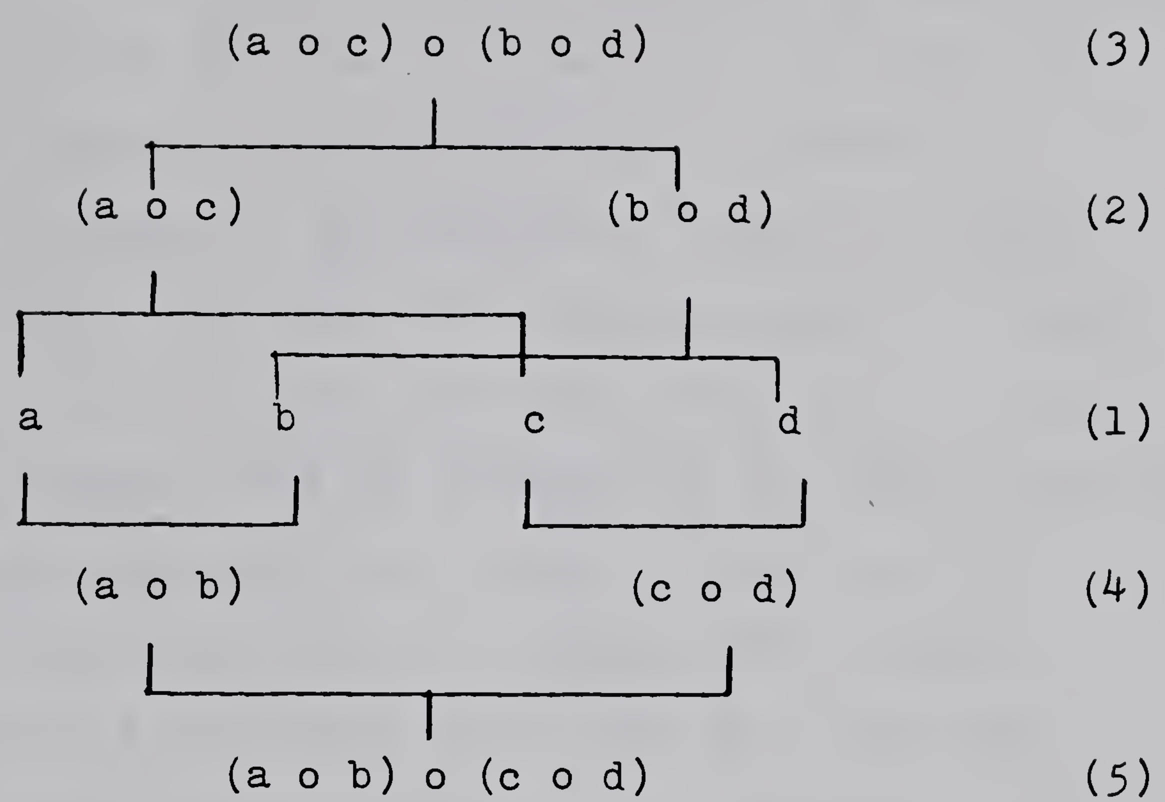


Figure 1: Schmeatic representation of the Bisymmetry Axiom. Starting at 1, and performing the bisections indicated by moving up or down the diagram, Bisymmetry hold if the bisection point given by 3 equals that given by 5.

illustrated by 1, 4, and 5 in Figure 1.

The intervals from nonadjacent standards, (a,c) and (b,d), may also be bisected to obtain bisection points (a o c) and (b o d). The interval defined by these bisection points may then be bisected to obtain (a o c) o (b o d). This latter sequence is illustrated by 1, 2 and 3 in Figure 1. According to the bisymmetry axiom, the final bisection points from these two sequences should be equal: that is, (a o b) o (c o d) = (a o c) o (b o d). It does not follow, however, that the distance of the final bisection points from points a and d are equal. That is, if the bisection points from the two sequences are x' and x", then the pairs of intervals (a,x') and (a,x") are not necessarily equal to (x',d) and (x",d). However, (a,x') does equal (a,x") if the bisymmetry axiom holds.

Fagot and Stewart (1970) have proposed an additional axiom that has been implicated in most previous attempts to validate scales constructed on the basis of the bisection operation. It is the following commutativity axiom.

6. Commutativity Axiom. For all a, b in P, (a o b) = (b o a).

They pointed out that if, in addition to Axioms 1 through 5, commutativity holds, then the following condition holds:

> Non-Parametric Scalability. For all a, b in P, (a o (a o b)) o ((a o b) o b) = (a o b).

Non-parametric scalability is presented schematically in Figure 2. Given any two standard stimuli a and b, the interval (a,b) may be bisected to obtain the bisection (a o b). This is represented in line 2, Figure 2. The intervals (a,(a o b)) and ((a o b),b) may be bisected to obtain bisection points (a o (a o b)) and ((a o b) o b). The interval defined by these last two bisection points may also be bisected to obtain (a o (a o b)) o ((a o b) o b). This sequence is illustrated by lines 3 and 4 in Figure 2. According to non-parametric scalability, the first bisection of the interval (a,b) should equal the last: that is, (a o b) should equal (a o (a o b)) o ((a o b) o b). It follows, from non-parametric scalability, that the first and last bisection points obtained from the two sequencies are equidistant from a and b. That is, if the bisection points from the two procedures are x' and x", then the intervals (a,x'), (a,x"), (x',b) and (x",b) are all equal.

Pfanzagl (1968) has shown that if commutativity holds, then $p = q = \frac{1}{2}$. Thus, Equation 2 reduces to:

$$F(a \circ b) = \tfrac{1}{2}F(a) + \tfrac{1}{2}F(b). \qquad (3)$$

This means that it is not necessary to estimate p in constructing the interval scale. For finite stimulus sets, failure of non-parametric scalability may be due to violations of either the reflexivity, commutativity or bisymmetry axioms. If failure of non-parametric scalability

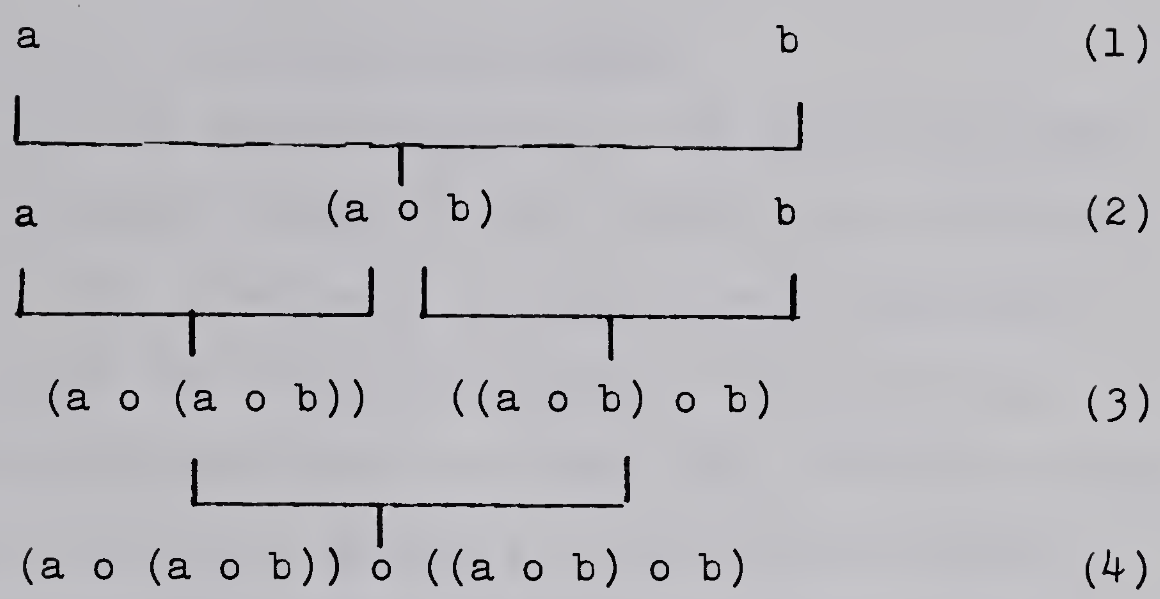


Figure 2: Schematic representation of Non-Parametric Scalability. Starting at 1, and performing the bisections by moving down the diagram, Non-Parametric Scalability holds if the bisection point given by 4 equals that given by 2.

is due to failure of commutativity, an interval scale representation may still be obtained as long as the bisymmetry axiom is not violated.

Empirical Tests of the Bisection Model

The numerous investigations of the bisection model have been concerned primarily with testing non-parametric scalability. There has been only one direct empirical investigation of the bisymmetry axiom. The previous investigations of the bisection model are reviewed in this section with particular emphasis on the methodological and theoretical issues that they have raised. For the most part, the method of adjustment has been used which can introduce biases in the results that are unique to the method.

Method of Adjustment. The first bisection experiment was devised and performed by Plateau in 1852, but the results were not published until 1872 (in Marks, 1974). Plateau presented a white and black surface to eight artists and asked them to paint a gray that appeared to be halfway between them. He found that the resulting grays were virtually identical even though they were painted under different illumination conditions. The reflectance of the bisection point was found to be above the geometric mean reflectance of the endpoints and Plateau concluded that the relation between the physical intensity of the

stimulus and the subjective response was a power function rather than a logarithmic function (Stevens, 1975).

One criticism of Plateau's results is that the white and black were presented side by side making it possible that contrast effects could have biased the responses. Another point to note is that Plateau's conclusion was based on a single bisection of a pair of stimuli so that his experiment cannot be considered as a test of any axioms of the bisection system.

Somewhat later, however, Gage (1934a, 1934b) investigated the possibility of constructing a scale of sensation for the loudness of tones and the brightness of lights by using the method of adjustment to test non-parametric scalability. In both experiments, subjects were presented with two stimuli with intensities a and b, and they were instructed to adjust a variable stimulus to bisect the interval. This procedure was repeated according to the sequence presented in Figure 2.

Gage found that the original bisection point, x', usually was not equal to the final bisection point, x", for either brightness or loudness. One exception was that when the endpoints were presented so that the brightest endpoint was to the subject's right, the original and final bisection points were approximately the same for one of his two observers. Gage concluded on the basis of his data that, in general, scales of sensation cannot be constructed for brightness and loudness using the method

of bisection. However, the discrepancy between x' and x" may have been due to the experimental operations used in scaling brightness and loudness. In the brightness study, the stimuli were presented simultaneously with the brighter endpoint to the subject's right on half of the trials. Gage found that when stimulus b, the brighter stimulus, was to the subject's right, both the original bisection point, x', and the final bisection point, x", were adjusted higher than when stimulus b was to the subject's left. He also found that, independent of the order in which the endpoint stimuli were presented, the bisection point x' was adjusted to be less intense than the bisection point x" (i.e., x" > x').

In the loudness study, the endpoint stimuli, a and b, were presented sequentially rather than simultaneously as in the brightness study, and the less intense endpoint stimulus was always presented first. In this study Gage again found that the bisection point x" was consistently greater than the bisection point x'. In neither study was the initial setting of the adjusted stimulus reported.

Unfortunately, there were shortcomings in both the design of Gage's experiments and his analysis of the data. For example, in the loudness study, the initial bisection point x' was determined by averaging the results from several bisections of a and b. Although the adjusted stimulus was continuously variable, the endpoint standards were calibrated in 5 decibel steps. Consequently, the

exact value of the obtained x' could not be presented as an endpoint stimulus. In an effort to overcome this difficulty, Gage used as endpoints two stimulus intensities, x_1' and x_2', which bracketed the obtained x'. Then to estimate the bisection point of the interval (a,x') say, subjects bisected the intervals (a,x_1') and (a,x_2') and the bisection point (a o x') was obtained by interpolation between the bisection points $(a \circ x_1')$ and $(a \circ x_2')$ using the method of proportional parts. Here lies the difficulty. The method of proportional parts makes the assumption that the relation between the stimulus intensity and the subjective response in the interval (x_1',x_2') is linear, which is clearly not the case. The decibel scale is logarithmic and even the sone scale for loudness is distinctly nonlinear on the physical scale. A similar situation existed for the brightess study because the filters used to determine the intensities of the stimuli were in discrete steps.

Additional difficulties with Gage's loudness study were pointed out by Newman, Volkmann and Stevens (1937). They noted that the endpoint stimuli were always presented in order of increasing intensity (ascending trials) and never in order of decreasing intensity (descending trials). According to Newman, et al., failure to balance ascending with descending orders leads to constant errors which must be taken into account before it can be concluded that valid sensory scales cannot be constructed using the method of bisection.

In a replication of Gage's loudness experiment in which ascending and descending trials were counterbalanced, Newman, et al. found x" to be only 0.26 and 0.18 decibels greater than x' for their two subjects. These differences are much smaller than Gage's values of 5.2 and 6.5 decibels based on ascending trials only. Newman, et al. also questioned the small range of loudnesses used by Gage and his selection of a range from stimuli of low intensity. Although Gage used a 38.3 decibel range of intensities (from 20.3 to 58.6) as compared with the 20 decibel range (from 80 to 100) used by Newman, et al., on the sone scale of loudness, Gage's range was approximately 0-7 sones while Newman, et al.'s range was approximately 30-70 sones. The latter investigators pointed out that if there are sources of constant error in the method of bisection, they would be magnified at low loudnesses and for small ranges. They maintained that the method of bisection yields results equivalent to other methods if sources of constant and variable error are minimized.

Stevens (1957) has called the systematic difference between bisection points obtained from ascending and descending trials hysteresis. He found that for loudness subjects set the adjustable stimulus 5 to 8 decibels higher on ascending trials than on descending trials. For brightness the difference was somewhat smaller. However, Stevens used a method of equisection in which the subjects

adjusted 3 or more variable stimuli between two endpoints until the intervals between adjacent stimulus pairs all appeared to be equal. He did not indicate the initial intensities of the variable stimuli used in his experiment.

Garner (1954) used both equisection and fractionation to construct a scale of loudness - the lambda scale. It was assumed that equisection yields a scale with equal intervals but an arbitrary origin, whereas fractionation provides a scale with a defined zero point but is based on an unknown subjective ratio. The method of fractionation required the subject to adjust a variable stimulus to some subjective ratio (usually one-half) of a standard stimulus. However, Garner argued that subjects may have difficulty in defining what combination of subjective intensities defines a ratio of one-half and, consequently, he was willing to assume only that the ratio was constant.

Garner controlled for another source of constant error in addition to controlling for hysteresis, the error due to the initial intensity of the variable stimuli. In the equisection task, the variable stimuli were set equal to the higher endpoint for half of the trials and for the other half they were set equal to the lower endpoint. In the fractionation task, the variable stimulus was set equal to the standard on half of the trials and to some lower intensity on the other half.

For equisection, Garner found that the bisecting intensity was set 1.5 decibels higher when the variable

stimuli were initially set to the higher endpoint intensity than when they were set at the lower. His subjects were instructed to listen to the stimuli in ascending and descending order equally often, but no check was made to ascertain whether subjects actually did this.

In the appendix to his article, Garner reported an additional experiment in which he controlled for hysteresis and counterbalanced the initial intensity settings of the adjustable stimuli. Both ascending and descending trials were used, and the intensities of the adjustable stimuli were set either high or low. He found that for ascending trials the bisection point was set 5.8 decibels higher than for descending trials. Although the data and statistical analysis were not reported, Garner stated than when the variable stimuli were initially set equal to the higher endpoint, the bisection point was adjusted 2 decibels higher than when the variable stimuli were set to the lower endpoint. This effect was independent of whether the trial was ascending or descending.

From Gage's (1934b) and Garner's (1954) results, it is clear that there are at least two independent sources of error associated with the use of the method of adjustment in the bisection task. The first is due to presenting the endpoints in ascending or descending order. The term hysteresis has been used to refer to these temporal effects of sequentially presented stimuli. (These temporal effects are interpretatively different from the left-right position

or spatial effects of simultaneously presented stimuli.) The second source of error is the initial intensities at which the variable stimuli are set in the method of adjustment.

There may be other sources of error as well. For example, in Garner's (1954) equisection task, subjects were instructed to determine the bisection point first, and then determine the quartersections. However, they were allowed to adjust all of the points again after the quartersections had been determined. It is not known whether or not bisection points obtained in an equisection experiment are equivalent to the bisection points obtained in the usual bisection experiment in which partitions of the interval are wholly sequential. Also, the method of adjustment does not prevent the subject from adjusting the variable stimuli to any value he desires at any time during a trial. This latter freedom could negate the experimental efforts to control for the error associated with the initial intnesity setting of the variable stimulus.

As mentioned previously, experimental efforts to validate bisection as a measurement system have, for the most part, tested non-parametric scalability. A single exception was provided by Cross (1965; Cross' experiment is also described in Coombs, Dawes & Tversky, 1970) in a direct test of the bisymmetry axiom using the method of adjustment to bisect loudness pairs. Cross controlled for hysteresis by presenting ascending trials half of the time and descend-

ing trials the other half, but he did not indicate whether he attempted to control for other sources of error.

Cross concluded that the bisymmetry axiom was satisfied to a good approximation, but the bisection point was somewhat higher for ascending trials than for descending trials. The mean difference between the final bisection points were .15 and -.45 decibels for ascending and descending trials respectively, for one subject. The ascending bisection points averaged 4.2 decibels higher than the descending bisection points. This is interesting because even though hysteresis was present, bisymmetry was satisfied suggesting that failure of non-parametric scalability may be due to violations of reflexivity or commutativity. One difficulty with Cross' findings is that his subjects performed all of the ascending bisections in the first half of the experiment and all descending bisections in the second half, instead of randomally presenting the two types of trials. It is not clear that learning, memory, temporal order, or differential adaptation effects are equivalent for the two different methods of trial presentation.

It can be seen that the bisymmetry axiom may hold even though non-parametric scalability does not hold, due to violations of either reflexivity or commutativity. From a review of the findings from bisection of loudness and pitch, Pfanzagl (1968) concluded that reflexivity is usually not violated but commutativity usually is, except for some persons having perfect pitch.

Pfanzagl also suggested that Gage's (1934a, 1934b) tests of non-parametric scalability do not support a violation of the bisymmetry axiom, but rather suggest a violation of commutativity. Recall that Gage's (1934a) loudness study used only ascending trials, and furthermore, his brightness study blocked ascending and descending trials in the same manner employed by Cross (1965). Thus, violations of commutativity reflect what is referred to as hysteresis, the temporal effects of presenting the stimuli sequentially, and it seems that if hysteresis is present, bisymmetry can hold even though non-parametric scalability is violated.

Fagot and Stewart (1970) made several criticisms of the Gage (1934a) and Newman, et al. (1937) studies on somewhat different grounds. First, they pointed out that these investigators assumed that non-parametric scalability should apply to the averages of several determinations of the bisection point, without justifying which of the numerous available measures of central tendency is appropriate. The result of the test may have been dependent upon which measure was actually used.

They also advanced a second criticism. Both the Gage and Newman, et al. studies used a single pair of endpoints of such limited range as to preclude general statements concerning the entire range of stimulation. Gage's interval was rather low on the subjective scale, and although the interval used by Newman, et al. was of

greater intensity and defined a subjectively wider range of stimulation, the entire range of sensitivity was not investigated. Fagot and Stewart argued that considerably more information can be obtained if several pairs of standards are used.

Finally, Fagot and Stewart interpreted p in Equation 2 as a response bias paramter and pointed out that neither Gage nor Newman, et al. had made an attempt to introduce this parameter to account for their data.

Consequently, Fagot & Stewart conducted a study using nine pairs of brightnesses as endpoints covering a range of stimulation from 0.1 to 600 Foot-lamberts. A trial consisted of a single bisection of the intervals represented by 1, 2 and 3 in Figure 2 for each pair of endpoints before going on to another trial. In analyzing their data, Fagot & Stewart divided the results into a construction set consisting of the three bisection points represented by 2 and 3 in Figure 2 and a test set consisting of the final bisection point represented by 4 in Figure 2. The construction set was used in estimating the parameters of a psychophysical function and the test set was used in a test of the predictions of non-parametric scalability.

A graph of the initial bisection of the interval (a,b), x', against the final bisection point, x", indicated a systematic bias such that x" tended to be overestimated relative to x'. The systematic bias may have been due to violations of reflexivity, commutativity or bisymmetry. To

investigate possible violations of reflexivity, Fagot and Stewart set $r = 0$ in Equation 1 and estimated $p + q$, which they found to be approximately equal to unity. They concluded that reflexivity was not violated and attribute the systematic bias to left dominance on the basis of some findings from an unpublished study. In that study, Fagot & Stewart found that subjects tended to judge the stimulus on the left as brighter over many replications of the pair in both left-right and right-left orders. This suggests that F(a) in Equation 1 is inflated and consequently p must be less than $\frac{1}{2}$ for that expression to hold. Also, the implication is that commutativity is volated and that x'' should be greater than x' which is consistent with their findings in the main experiment. Fagot and Stewart concluded that neither their study nor Gage's (1934b) brightness study provided evidence to suggest that bisymmetry is violated.

Since non-parametric scalability did not hold, it was not possible to estimate the value of the bias parameter p unless the psychophysical function can be specified. The psychophysical function relating perceived intensity to the measure of the physical stimulus is most frequently assumed to be a power function of the form:

$$R = aX^n,$$

where R is the response, X the physical measure of the stimulus and a is a constant associated with the units of measurement of the stimulus values. The exponent n is an

index of the sensitivity of the modality under study.

Using two variations of this simple power law, one with a correction on the stimulus axis and the other with a correction on the response axis, Fagot and Stewart obtained estimates of p less than $\frac{1}{2}$ in all but one instance. They concluded that this supported their hypothesized left dominance bias as an explanation of the failure of non-parametric scalability. The estimates of n were small and negative, however, which is not consistent with previous findings of n equal to approximately .33 for brightness, nor does it have a psychological interpretation.

A graph of the predicted values of x' obtained from the fit of Equation 2 to the data against the obtained values of x" indicates that the systematic biases evidenced earlier were substantially reduced. However, the predicted values were still too low at the low end and too high at the high end. This does not imply that bisymmetry is violated since non-parametric scalability requires reflexivity and commutativity to hold before it can be satisfied. Further, the remaining deviations cannot necessarily be attributed to violations of the bisymmetry axiom because they are completely confounded with any inaccuracies that may exist in the assumed form of the psychophysical function used in estimating the parameters. In fact, Fagot and Stewart attribute the systematic deviations to this latter source.

Other researchers, who did not directly test the axioms of the bisection system, have concluded that valid

sensory scales can be constructed from the method of bisection. For example, Weiss (1975) had subjects adjust Munsell grays to obtain bisections, trisections, and quadrisections of pairs of endpoint stimuli and he applied functional measurement theory to the data from the equisection tasks on the assumption that equisection obeys an averaging model. If the data are additive or if a polynomial transformation can be found that renders the data additive, the model is supported. Furthermore, the form of the transformation defines the psychophysical function. Weiss found that the linear model applied to the data from bisection, but not to the data from the other procedures. He concluded that the scale constructed by the bisection method was consistent with a scale of this attribute obtained by applying the functional measurement approach to judgments of average grayness (Weiss, 1972).

Several comments can be made regarding Weiss' study. Although a specified form of the psychophysical function was not assumed as in the Fagot and Stewart (1970) study, the transformation necessary to make the data additive is required and that transformation is assumed to be linearly related to the psychophysical function. In addition, while Weiss did not specifically report whether he used trials with the endpoint stimuli in reversed spatial order, it appears that he did not. This is unfortunate because that information might have shed some light on the left dominance hypothesis of Fagot and Stewart.

<u>Method of Constant Stimuli.</u> A common characteristic of the foregoing studies in that the method of adjustment was used in the production of the bisecting stimuli. In using the method of adjustment it is assumed that subjects are capable of matching intervals, that is, that they can adjust a variable stimulus such that it is equidistant from each of the two endpoint stimuli.

An alternative strategy, employed by Beck and Shaw (1968), is the use of the method of constant stimuli. In this method, the subject is presented with two endpoint stimuli and a variable stimulus. He is required to judge whether the variable stimulus is more similar to the lower or upper of the two endpoint stimuli. A set of variable stimuli are selected so that the bisection point is highly likely to fall within its range. The frequency of times that each variable stimulus is judged to be more similar to the higher endpoint simulus should provide a basis for the construction of a psychometric function. The bisection point may then be defined as that physical value of a stimulus for which the relative frequency of judgments that it is more similar to the higher endpoint is .5.

Beck and Shaw (1968) investigated loudness bisection using the method of constant stimuli. On each trial, the subject was presented with a series of tones, and his task was to judge whether the variable tone was more similar to the lower or higher intensity endpoint (standard). The endpoint stimuli were pairs of 1000 Hz tones separated by

15 decibels and the variable stimulus sets consisted of five tones of the same frequency within the range defined by the endpoint stimuli. After transforming the frequency data into standard normal scores, Beck and Shaw determined the stimulus intensity for a proportion of .5 (i.e., $z = 0$) by visually fitting a line to the data for each endpoint pair. They found that the bisection point was close to the mean of the endpoint pairs in decibels. Beck and Shaw reported their results to be similar to those of Garner (1954) who used the method of adjustment, and they viewed their task as essentially a bisection task with a different methodology.

Beck and Shaw controlled for hysteresis by pooling frequencies from ascending and descending trials. Separate analyses of the two sets of data revealed that ascending trials yielded a higher bisection point than descending trials, and this result decreased as absolute stimulus intensity increased. This is consistent with findings presented earlier (Garner, 1954). However, Beck and Shaw did not attempt to test either non-parametric scalability or the bisymmetry axiom.

The method of constant stimuli has two advantages over the method of adjustment for a test of the bisection model. First, the task of judging whether the variable stimulus is above or below the midpoint of the interval defined by the endpoint stimuli appears to be an easier task for subjects than the direct determination of the bisection point in the method of adjustment. This is

particularly the case for auditory or visual stimuli. The method of constant stimuli also allows for investigation of the bisection model for stimulus continua that preclude easy production of a stimulus (e.g., lifted weight, taste, odor).

Second, the method of constant stimuli allows for the collection of a second kind of data, the latency of the judgmental response. These data may constitute a second operation for the evaluation of the bisection experiment, one which should converge with the results from the frequency data. The rationale for the use of reaction time data in the context of the bisection task is based on previous evidence that the time required for judging which of a pair of stimuli has the greater magnitude is related to the difference between the subjective magnitudes of the stimuli by a reciprocal function (Curtis, Paulos & Rule, 1973; Mullin & Curtis, 1973).

In applying the method of constant stimuli to the bisection task, the choice task is one in which subjective intervals are compared. That is, for endpoint stimuli a,b and variable stimulus i, it is assumed that the subject's task is to evaluate which of the two subjective intervals, (a,i) or (i,b), is the larger. It is assumed that the larger the difference between the two subjective intervals, the shorter should be the latency of response. This relation is given by the expression:

$$RT_{abi} = f\left[d(a,i) - d(i,b)\right]^{m} + b\,, \qquad (4)$$

where RT represents the choice reaction time for choosing between the subjective intervals d(a,i) and d(i,b), m is an output or response parameter which should be negative, and b is an additive constant reflecting what is known as the irreducible minimum response time.

A considerable body of evidence from direct judgments of differences (Curtis, et al., 1968; Rule, Curtis & Markley, 1970) and choice reaction time (Curtis, Paulos & Rule, 1973; Mullin & Curtis, 1973) suggest that the subjective difference between two stimuli is equal to the difference between their subjective magnitudes. Consequently, Equation 4 may be rewritten in the form:

$$RT_{abi} = f\left[(Y_i - Y_a) - (Y_b - Y_i)\right]^m + b, \quad (5)$$

where Y's are the subjective magnitudes of the endpoint and variable stimuli a, b and i, respectively. Further if the subjective magnitude of a stimulus is related to its physical measure by a power function of the form $Y = aX^k$, where k is a sensory or input parameter, then Equation 5 becomes:

$$RT_{abi} = a\left[(X_i^k - X_a^k) - (X_b^k - X_i^k)\right]^m + b. \quad (6)$$

This expression simplifies to:

$$RT_{abi} = a\left[\left|X_a^k + X_b^k - 2X_i^k\right|\right]^m + b. \quad (7)$$

The absolute value signs have been added because it is assumed that response time is a function only of the difference in magnitudes of the intervals and not of the algebraic sign of the difference. A solution for the parameters a, k, m and b in Equation 7 should make it possible to predict the bisection points from the reaction time data, allowing for a test of the bisymmetry axiom from these data as well as a verification of the results obtained from the frequency data.

METHOD

Subjects

The subjects were four graduate students in psychology who were paid for their participation. Three of the subjects participated in 15 sessions while one subject participated in 35 sessions. Each session was approximately one hour in length.

Apparatus

Subjects were tested in a black booth with their heads resting on a chin rest. Mounted in front of the subject was a box containing three buttons: one button initiated the trial sequence and the other two buttons were for responding. The stimuli, presented in Maxwellian view through a 2-mm artifical pupil, were produced by a Monsanto Light Emitting Diode (LED) MV 5752, with a peak wavelength of 635 nm (red). The stimulus subtended approximately 1.75 degrees of visual angle. The presentation sequence consisted of depression of the start button, a 600 millisecond delay, followed by three electronically controlled, sequentially presented stimuli of 700 milliseconds duration each, with interpulse intervals of 600 milliseconds. The onset of the second stimulus activated a Fluke 1950A Digital Counter which had been converted into an electronic timer. The counter was stopped by the

depression of either of the two response buttons. The stimulus intensities were controlled by potentiometers which were preset to the voltages required for the various stimulus intensities.

Stimuli

Standard stimuli for the entire stimulus range, endpoints a and d, were selected so as to cover as large a range of intensities as was possible given the capabilities of the apparatus. The intermediate endpoints, b and c, were selected so as to be approximately equally spaced between a and d on a scale of cd/m^2 when transformed to the 1/3rd power based on Stevens (1975) power function exponent of approximately .33 for brightness. That is, if the endpoints are raised to the 1/3rd power, the intervals between adjacent pairs are approximately equal. The four endpoints a, b, c, and d were 1.196, 11.567, 48.2 and 121 cd/m^2 respectively. Nine variable stimuli were associated with each pair of standard stimuli serving as endpoints. The variable stimuli for the initial four bisections are presented in Table 1. These variable stimuli were also approximately equally spaced on a scale of intensity raised to the 1/3rd power and they extended over the middle 50% of the range between endpoints after the power transformation. That is, if two endpoints, say a and b, along with their variable stimuli are raised to the 1/3rd power, the first variable stimulus is at approximately the 25% point, the

TABLE 1

Stimulus values for the initial four bisections in cd/m^2. Endpoints were selected to cover a large range of stimulation on a cube root scale. Variable stimuli were selected to cover the middle 50% of the range between the endpoints on a cube root scale.

	Stimulus Values		Stimulus Values
Endpoint a	1.196	Endpoint a	1.196
Variable 1	2.896	Variable 1	5.955
2	3.201	2	7.421
3	3.716	3	9.066
4	4.115	4	10.830
5	4.607	5	12.683
6	5.045	6	14.554
7	5.575	7	16.339
8	6.031	8	19.194
9	6.566	9	20.740
Endpoint b	11.567	Endpoint c	48.200
Endpoint c	48.200	Endpoint b	11.567
Variable 1	66.780	Variable 1	29.875
2	69.603	2	34.552
3	72.900	3	39.100
4	76.234	4	43.940
5	79.148	5	49.325
6	82.155	6	53.760
7	85.670	7	58.900
8	88.850	8	63.940
9	92.540	9	69.325
Endpoint d	121.000	Endpoint d	121.000

last at approximately the 75% point, and the intervals between adjacent variable stimuli are approximately equal. The final two sets of endpoints needed for testing the bisymmetry axiom depended on the outcome of these initial four bisections.

Once the endpoints for the final test bisections were determined for each subject, the variable stimuli were selected in the same manner as those for the initial four bisections. These are given in Table 2 for three subjects and in Table 3, Part a for the fourth subject. The stimuli in Part b of Table 3 were used in a replication of the experiment for that subject.

Between trials and during interpulse intervals, the LED was set to a low intensity of .29 cd/m^2 to serve as a fixation target for the subject.

Procedure

The experimental sessions can be divided into two phases, which are illustrated in Figure 3. In the initial phase, four bisections were obtained from each subject. This was followed by a test phase in which two final bisections were obtained. As indicated in Figure 3, the endpoints from the the final two bisections were based on the outcome of the bisections in the initial phase. Fifteen experimental sessions were required to complete the two phases, ten for the four bisections in the initial phase and five for the two test bisections.

TABLE 2

Stimulus values for the final test bisections in cd/m^2 for three subjects. Endpoints were estimated from the data and variable stimuli were selected to cover the middle 50% of the range between the endpoints on a cube root scale.

	Stimulus Values		
Subject	DM	JE	KM
Endpoint (a o b)	5.53	5.19	2.72
Variable 1	16.54	19.19	8.58
2	18.24	21.51	9.62
3	20.63	23.83	10.92
4	24.55	26.18	12.33
5	25.74	28.84	13.68
6	28.50	31.53	15.26
7	31.66	34.48	16.94
8	34.48	37.85	20.85
9	38.48	40.95	20.85
Endpoint (c o d)	76.43	76.40	42.98
Endpoint (a o c)	15.70	12.55	5.24
Variable 1	22.06	20.30	9.06
2	23.16	21.40	9.62
3	24.18	22.72	10.28
4	25.19	23.95	11.02
5	25.96	25.52	11.84
6	27.07	26.73	12.42
7	28.15	28.15	13.31
8	29.42	29.53	13.96
9	30.34	31.29	14.76
Endpoint (b o d)	40.81	45.94	22.95

TABLE 3

Stimulus values for the final test bisections in cd/m^2 for subject LM. Endpoints were estimated from the data and variable stimuli were selected to cover the middle 50% of the range between the endpoints on a cube root scale. LM1 gives stimulus values for the final test bisections for the basic experiment and LMP gives stimulus values for the final test bisections for the replication.

Subject		Stimulus Values LM1	Stimulus Values LMP
Endpoint (a o b)		5.47	5.40
Variable	1	17.04	16.84
	2	19.19	18.77
	3	21.73	21.29
	4	24.40	23.37
	5	27.07	26.07
	6	31.02	29.19
	7	33.62	32.40
	8	37.37	36.00
	9	41.08	39.82
Endpoint (c o d)		82.81	83.26
Endpoint (a o c)		16.80	16.27
Variable	1	25.74	24.97
	2	26.85	26.07
	3	28.15	27.31
	4	29.53	28.61
	5	31.05	30.07
	6	32.53	31.49
	7	33.98	33.03
	8	35.35	34.55
	9	37.37	36.13
Endpoint (b o d)		52.48	52.70

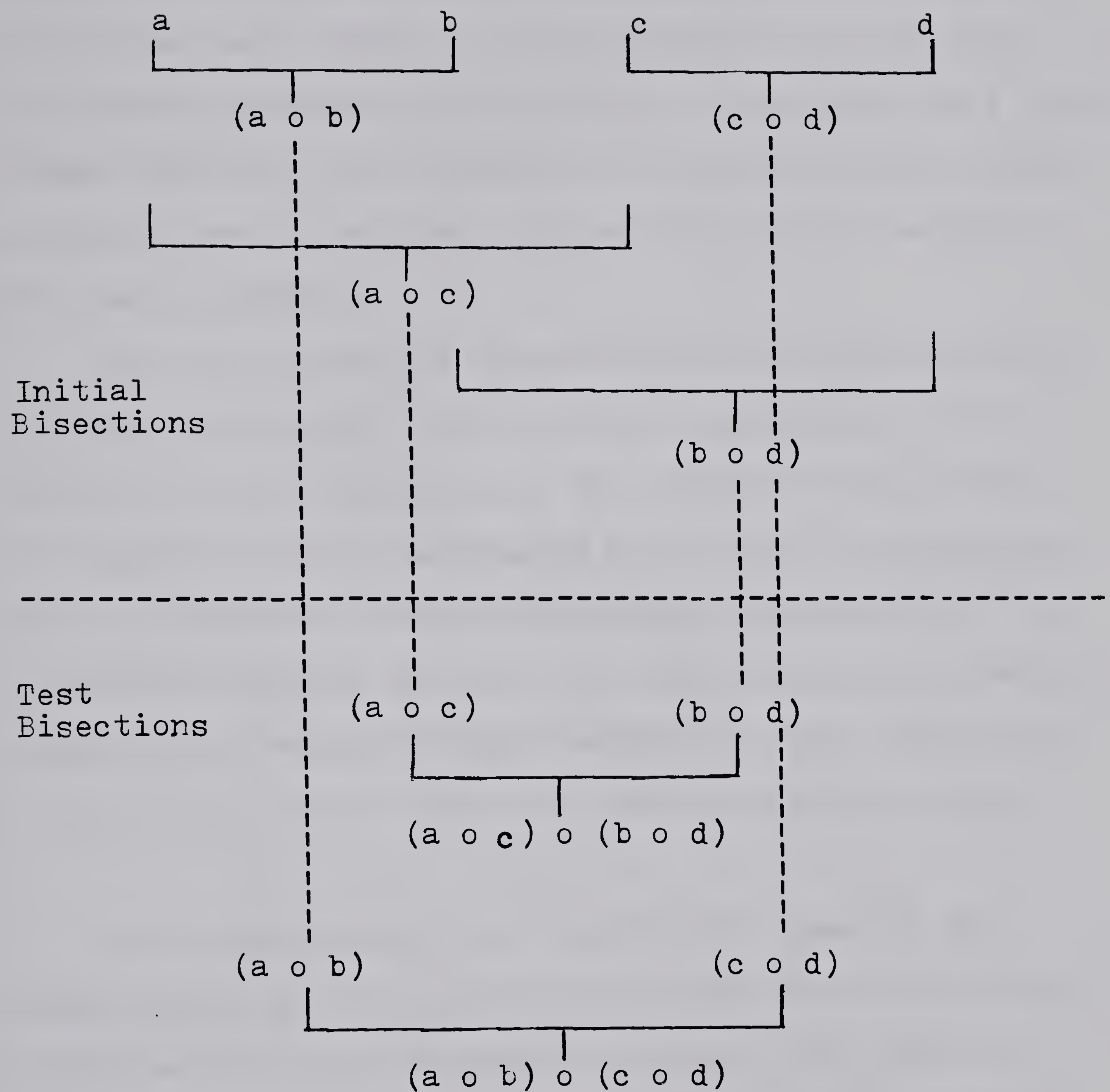


Figure 3: Schematic representation of the sequence of initial and test bisections obtained from each subject.

Two sets of endpoints were presented in each experimental session with the restriction that no stimulus serving as an endpoint in one set appeared as an endpoint in the other set. That is, sets of endpoints (a,b) and (c,d) may be presented in a session, but not (a,b) and (a,c) because both sets share stimulus a. Within the two phases, all sets of endpoints were counterbalanced over sessions and between subjects.

Within a session, a subject received five blocks of 18 trials for each of the two sets of endpoints. The 18 trials in a block consisted of two presentations of the nine variable stimuli associated with a pair of endpoints, once in combination with an ascending presentation of the two endpoint stimuli and once in a descending presentation. A total of 25 responses were obtained for each combination of endpoints, variable stimulus and presentation order.

At the beginning of an experimental session the subject adjusted the chin rest and chair to be comfortable and dark adapted for 15 minutes. During that time, the subject was given the instructions in Appendix A. After dark adaptation, the experimenter sounded a tone to indicate that the subject could initiate the trial. The subject then pressed the start button, viewed the stimulus sequence and responded by pressing one of the response buttons as quickly as possible. The experimenter recorded whether the subject judged the variable stimulus to be

above or below the midpoint, as well as the latency of that response. The stimuli for the next trial were then set, and the tone was sounded to indicate that the next trial could be initiated. The first 8 trials were practice trials and were excluded from the data analysis.

After the presentation of 5 blocks for one pair of endpoints, there was a break, followed by the presentation of 5 blocks for another pair of endpoints.

In order to evaluate the reliability of the responses based on 25 replications, additional data was collected from one subject. For subject LM, 20 additional sessions were conducted. For the first ten of these, stimulus values were the same as for the previous initial ten sessions. These two sets of data were combined to determine endpoint values for the final test phase. The variable stimuli were selected in the same manner as those for the basic experiment. These stimuli are given in Part b of Table 3. This resulted in 50 replications in ascending and descending order for all pairs of endpoints and all variable stimuli. Ten additional sessions were required for the test phase. Owing to the different stimulus values, the data from the test phase of the basic experiment could not be combined with the data from the replication (See Table 3, Part a versus Part b).

RESULTS

Both frequency and reaction time data were obtained from each subject. The frequency data, which is discussed first, consisted of the frequency with which each variable stimulus was judged more similar to the brighter of the two endpoints, while the reaction time data are the latency between the onset of the variable stimulus and the response.

Frequency Data

The relative frequency that each variable stimulus was judged to be more similar to the brighter of the two endpoints was transformed to a normal deviate (z scores). Figures 4-9 present values of the normal deviate as a function of the variable stimuli for each pair of endpoints. Separate plots are presented for ascending and descending trials and for frequencies pooled over ascending and descending trials for each subject. The different symbols represent the data for the variable stimuli associated with different pairs of endpoints as indicated in the figure captions. Data for the subject who received 50 replications (LM) are presented in three figures. Those from the first 25 are in Figure 6, those for the four initial bisections in the replication are in Figure 7 and the data from all 50 replications are presented in Figure 8. These three sets of data are labeled throughout as LM1 for the initial

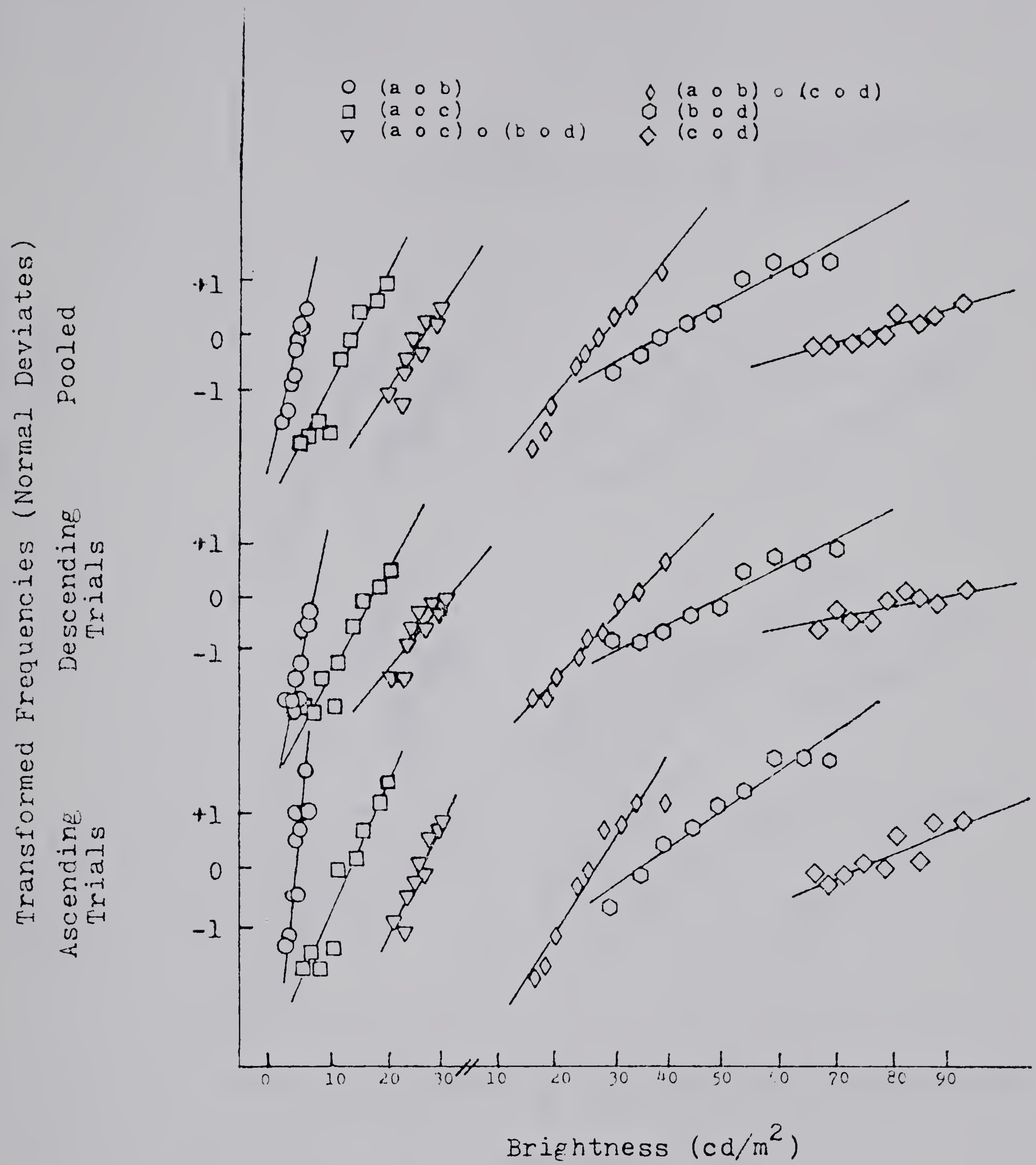


Figure 4: Transofmred frequencies (normal deviates) as a function of brightness for subject DM. Lines represent best fitting (in a least squares sense) linear equations for the variabile stimuli associated with each pair of endpoints.

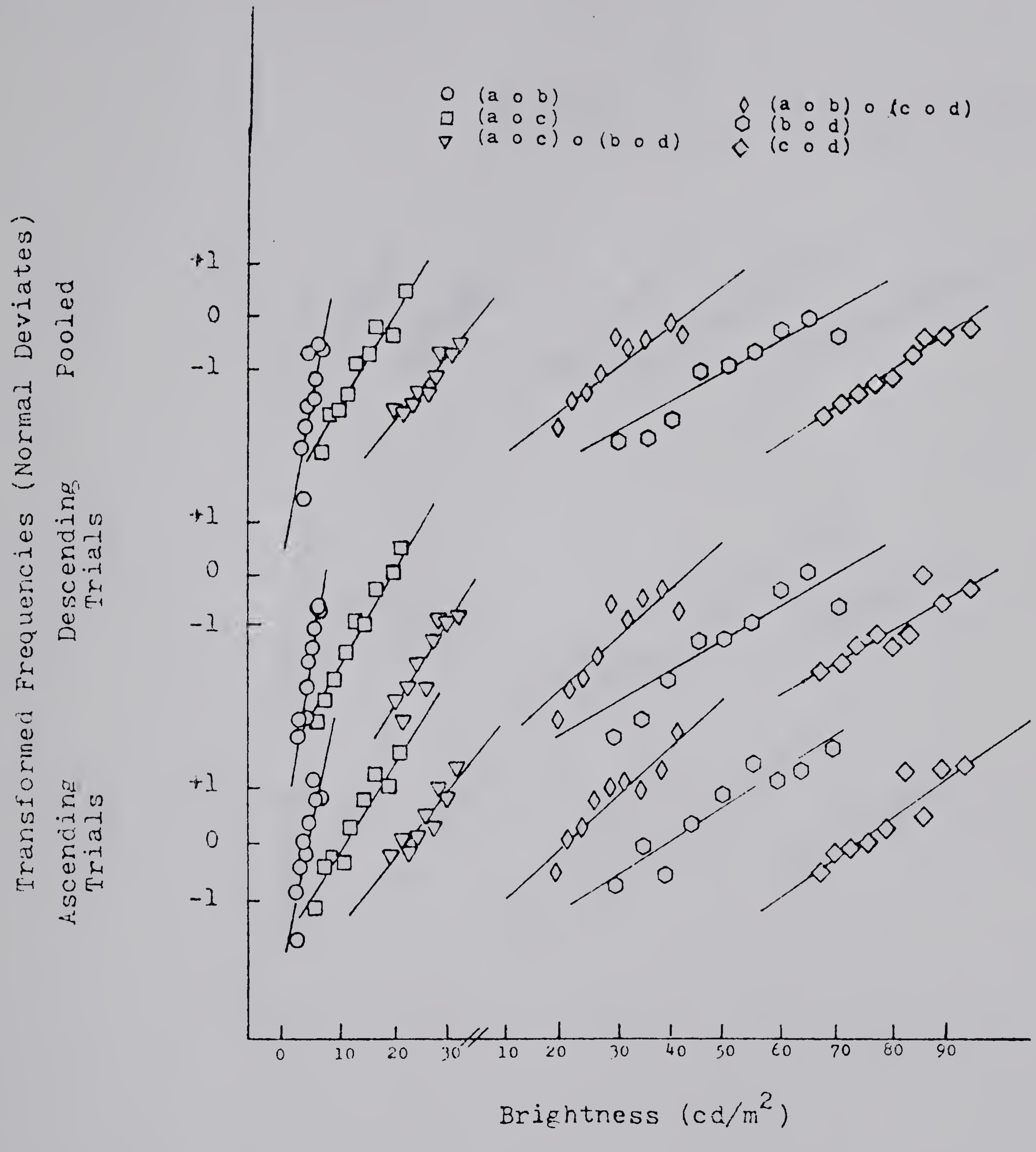


Figure 5: Transformed frequencies (normal deviates) as a function of brightness for subject JE. Lines represent best fitting (in a least squares sense) linear equations for the variable stimuli associated with each pair of endpoints.

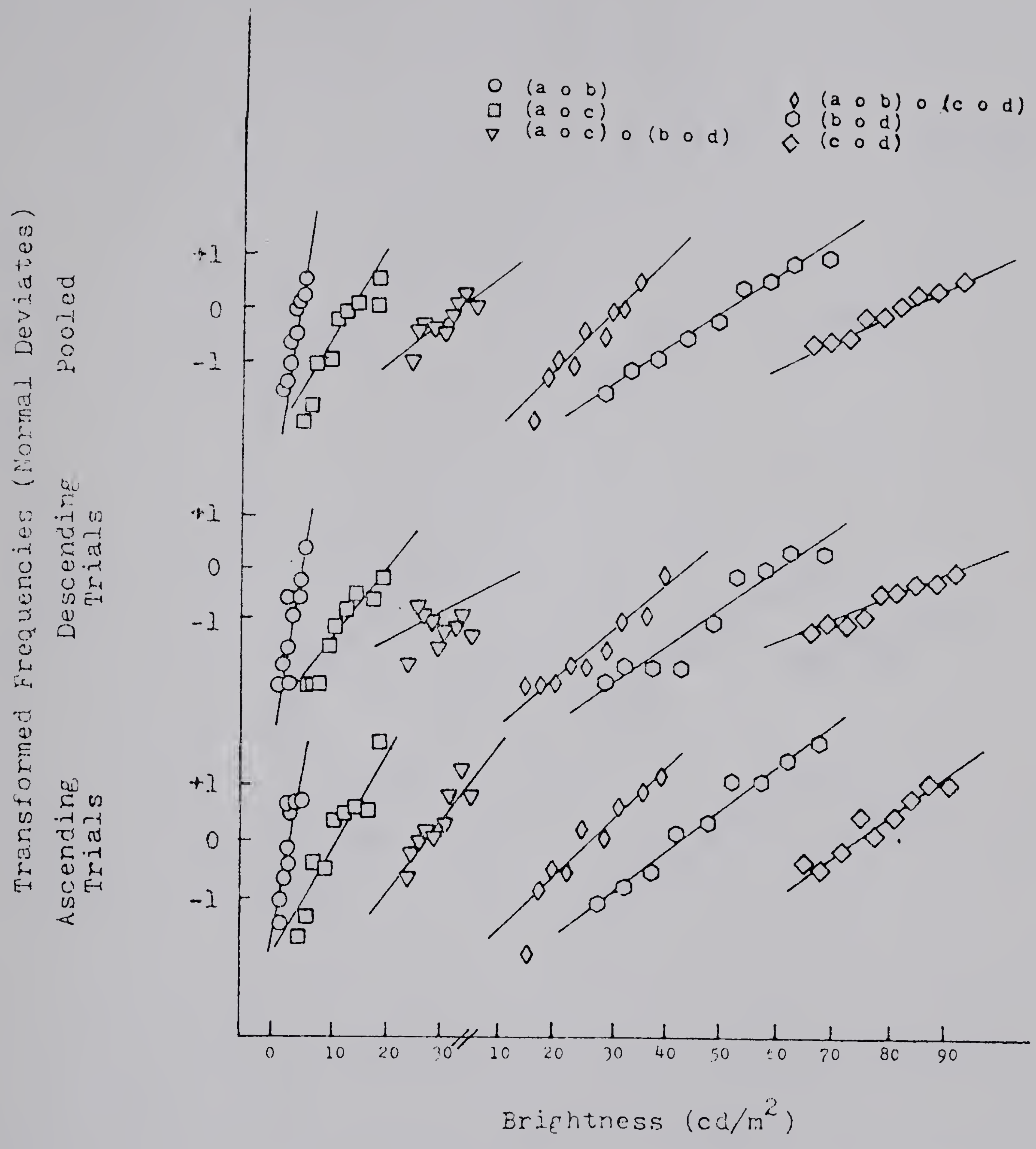


Figure 6: Transformed frequencies (normal deviates) as a function of brightness for subject LM1 (i.e., the basic experiment). Lines represent best fitting (in a least squares sense) linear equations for the variable stimuli associated with each pair of endpoints.

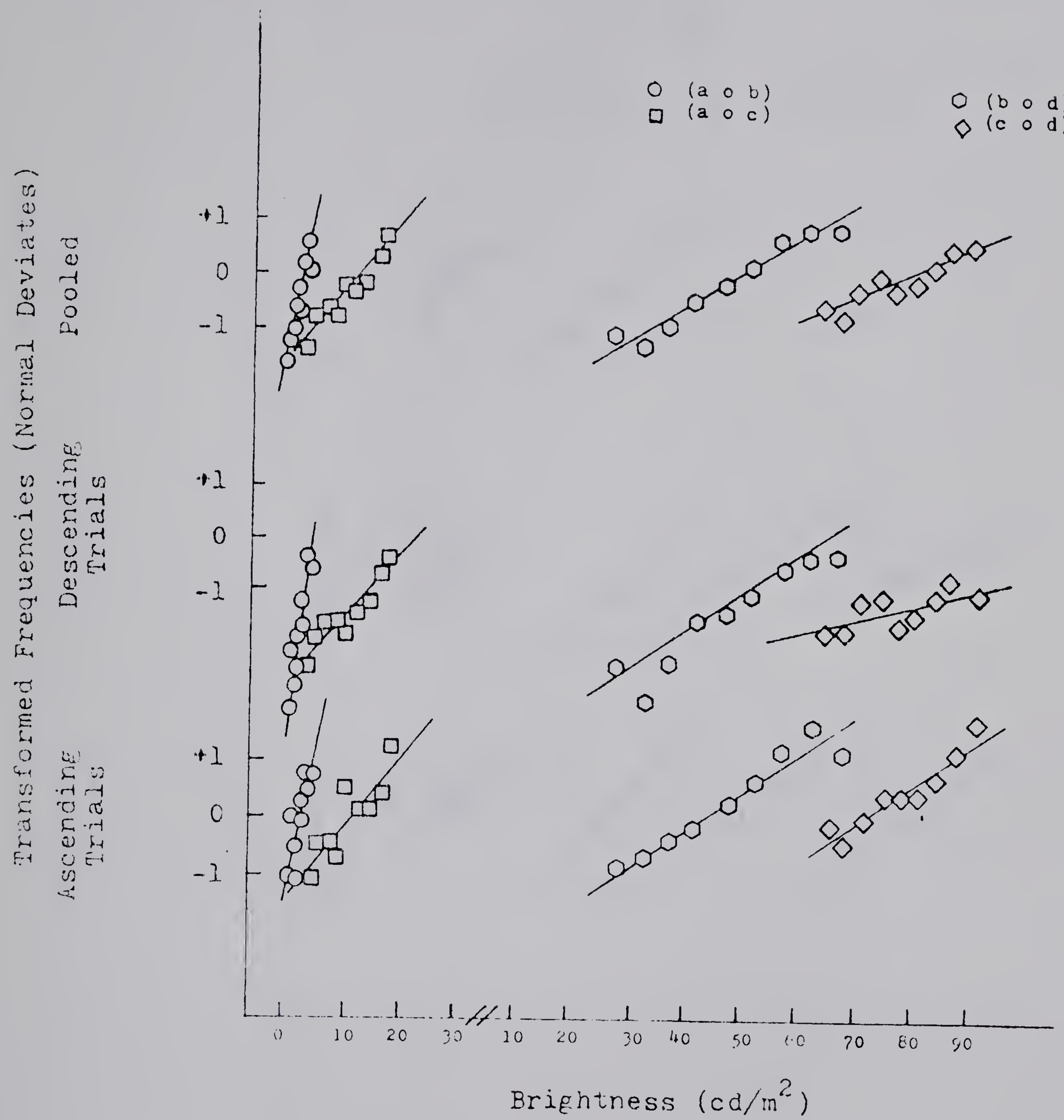


Figure 7: Transformed frequencies (normal deviates) as a function of brightness for subject LM2 (i.e, the first part of the replication). Lines represent best fitting (in a least squares sense) linear equations for the variable stimuli associated with each pair of endpoints.

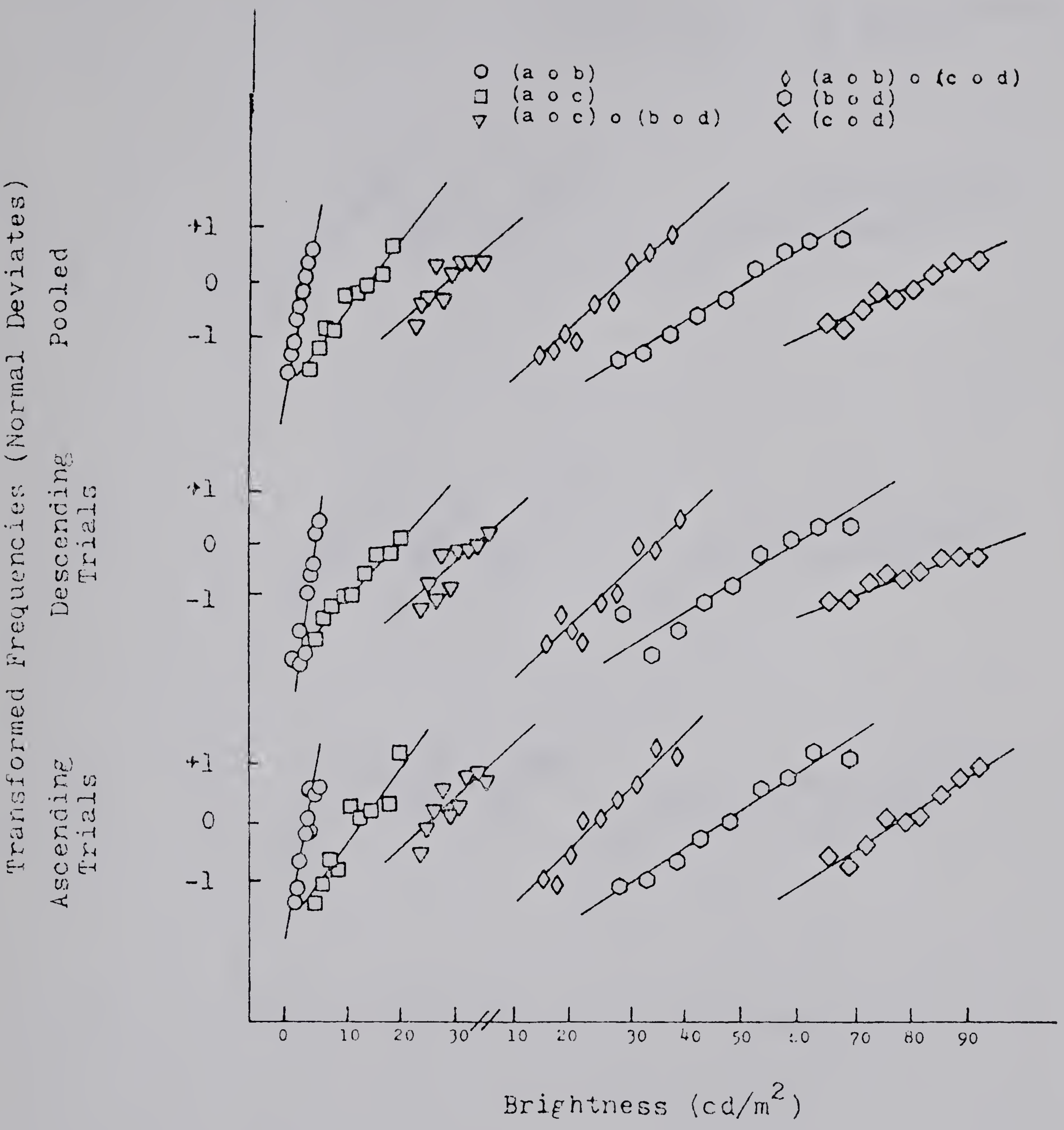


Figure 8: Transformed frequencies (normal deviates) as a function of brightness for subject LMP. The plots represent the data of Figure 7 combined with the comparable data of Figure 6, along with the data from the last part of the replication (the final bisections). Lines represent best fitting (in a least squares sense) linear equations for the variable stimuli associated with each pair of endpoints.

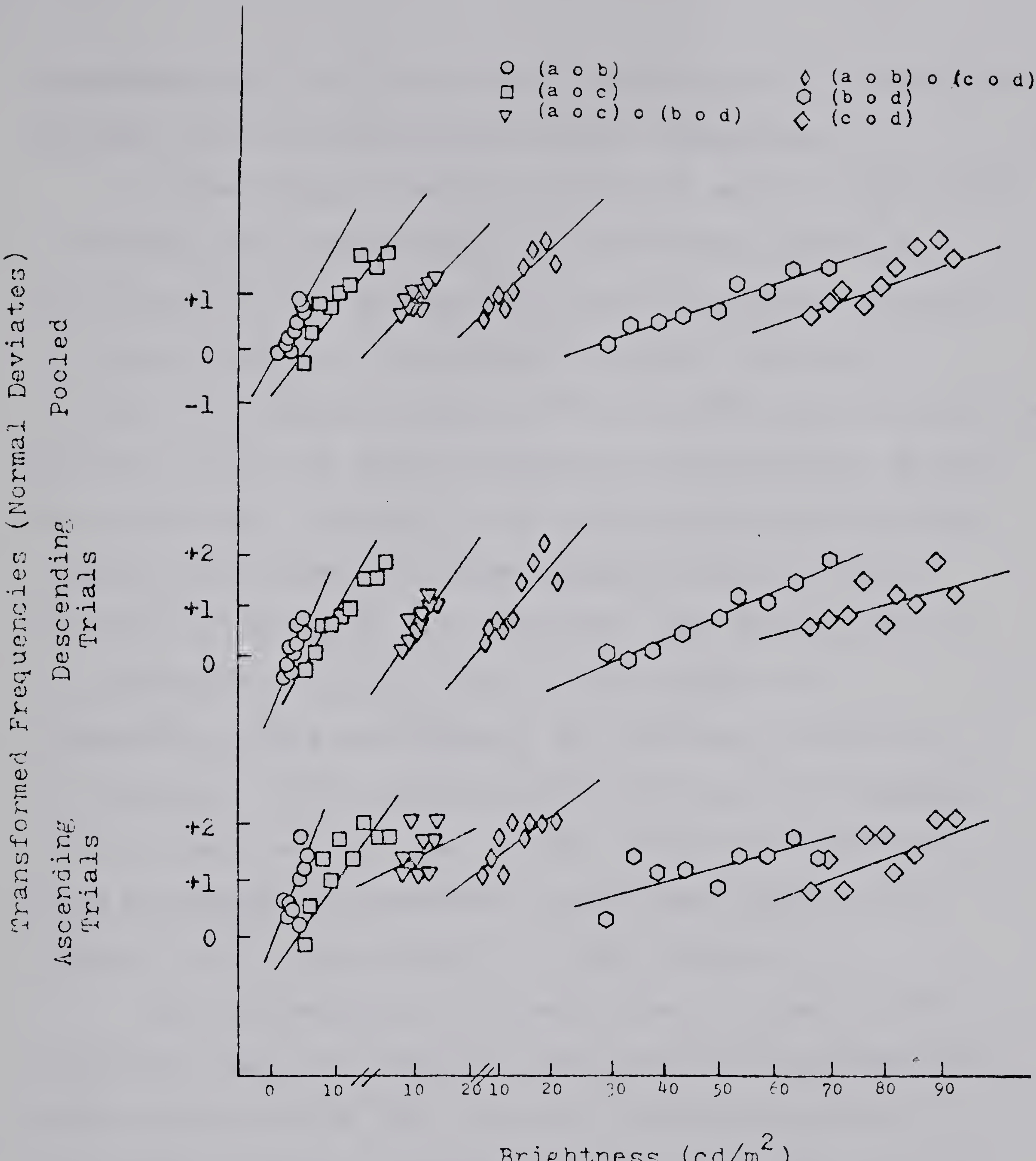


Figure 9: Transformed frequencies (normal deviates) as a function of brightness for subject KM. Lines represent best fitting (in a least squares sense) linear equations for the variable stimuli associated with each pair of endpoints.

experiment, LM2 for the initial bisections of the replication and LMP for the corresponding pooled frequencies.

A clear linear trend is evident in most of these plots, consistent with expectations that the normal ogive is descriptive of the psychometric function relating frequency to the intensity of the variable stimuli. Figures 4, 5, 7, and 8 for subjects DM, JE, LM2, and LMP show this trend very strongly with only minor violations of monotonicity evident in a few plots. Figures 6 and 9 which present the results from LM1 and KM show the same general trend but exhibit serious violations in two instances. The frequencies for the bisection ((a o c) o (b o d)) in Figure 6 for descending trials and Figure 9 for ascending trials do not increase as the variable stimulus intensity increases, and, in fact, show no clear trend. When the frequencies from ascending and descending trials are combined, the linear trend is more clearly evident, however.

When the method of constant stimuli is used in the bisection task, the point of bisection may be estimated as that stimulus value that is judged to be more similar to the higher standard 50 percent of the time. When relative frequencies are transformed to their corresponding normal deviates, as in Figures 4-9, the bisection point would be that stimulus value whose physical measure is associated with $z = 0$. In practice, of course, it would be highly unlikely that a stimulus of precisely this value would be represented within the set of variable stimuli. Consequently,

its value must necessarily be estimated from the empirical psychometric function. If it can be assumed that the frequency of times the variable stimulus is judged to be more similar to the higher standard increases with increasing intensities of the variable stimuli according to the cumulative normal distribution, the relation between the transformed relative frequencies and their associated variable stimulus measures should be linear. The best fitting linear function provides a reliable basis for estimating the value of the bisecting stimulus.

A linear least squares analysis between the variable stimuli and the obtained frequencies transformed into standard normal deviates was done for each set of endpoints using a Mueller-Urban weighting. The expression fitted was

$$z_i = aX_i + b, \tag{8}$$

where z's denote the normal deviates associated with the frequencies for each variable stimulus, X_i. The slope and intercept of the line are represented by a and b respectively. The procedure minimizes the weighted sum of the squared deviations of the normal deviates from the predicted line. The Mueller-Urban weights take into account the fact that chance errors in proportions near zero and one have a large effect on the corresponding standard normal deviate. Consequently, those proportions should not receive the same weighting in the least squares analysis as proportions near .5 where chance errors do not have much effect on the

standard normal deviates. Frequencies of zero and n were changed by $\pm$ 0.5 so as to base each solution on all the variable stimulus values (Woodworth & Schlosberg, 1954). This correction was not involved in any of the crucial analyses (i.e., those that determined the stimuli to be used as standards in subsequent bisections).

The least squares analysis was applied to transformed frequencies pooled over ascending and descending trials and to ascending and descending trials separately. The obtained least squares regression lines are included in Figures 4-9. In general, the fits of the lines to the data are quite good. The exception is again for the bisection ((a o c) o (b o d)) in Figure 6 (LM1) for descending trials and in Figure 9 (KM) for ascending trials. But, when the frequencies from those trials were pooled with the frequencies from corresponding ascending-descending trials, the fit was quite good.

Table 4 gives the proportion of variance accounted for by the linear regression. Except for KM, the proportion of variance accounted for is reasonably high in the majority of instances. This is to be expected when there is a monotone relation between stimulus intensity and response measures. Notice the small value for the bisection ((a o c) o (b o d)) for LM1 descending and KM ascending. This corresponds with the fact that the frequencies do not increase monotonically as stimulus intensity increases in these two instances.

TABLE 4

The proportion of variance accounted for by the least squares analysis between the transformed frequencies and the variable stimulus values for ascending descending and pooled trials.

Subjects		DM	JE	KM	LM1	LM2	LMP
Endpoints							
(a o b)	Ascending	.923	.827	.688	.869	.829	.918
	Descending	.867	.939	.770	.920	.861	.956
	Pooled	.953	.932	.827	.950	.935	.958
(c o d)	Ascending	.776	.838	.466	.910	.874	.920
	Descending	.682	.810	.435	.942	.502	.931
	Pooled	.888	.969	.566	.970	.853	.844
(a o c)	Ascending	.917	.899	.709	.830	.726	.839
	Descending	.930	.887	.938	.909	.841	.971
	Pooled	.942	.935	.851	.854	.903	.826
(b o d)	Ascending	.908	.851	.502	.970	.939	.968
	Descending	.926	.771	.936	.908	.917	.936
	Pooled	.956	.883	.944	.966	.952	.895
(a o b) o (c o d)							
	Ascending	.947	.868	.604	.880		.931
	Descending	.982	.832	.765	.913		.897
	Pooled	.968	.916	.806	.927		.887
(a o c) o (b o d)							
	Ascending	.908	.895	.086	.863		.719
	Descending	.834	.813	.708	.304		.772
	Pooled	.867	.935	.831	.748		.393

Estimated Bisection Points

The bisection point, defined as that stimulus value which is judged more similar to the higher standard 50 percent of the time has a standard normal deviate of zero. Consequently, the bisecting stimulus may be estimated by setting the left half of Equation 8 equal to zero and solving for the value of X. Estimates of the bisection points based on the frequencies pooled over ascending and descending trials are presented in Table 5 for each of the intervals bisected by each of the four subjects. Because the final two test bisections are based on endpoint standards defined by previous bisections by the same subject, it is not appropriate to pool results for these bisections over subjects.

There were no proportions of zero or one in the frequencies for the initial four bisections. Thus, the correction of $\pm$ 0.5 mentioned earlier did not affect the analyses that determined any of the bisection points that were subsequently used as endpoint standards.

An inspection of Table 5 reveals that the results from subject KM are not only anomalous, but they lack sufficient internal consistency to provide a basis for testing the bisymmetry axiom. In particular, the bisection of the intervals (2.72,42.98) and (5.24,22.95) cd/m^2 yielded bisection points of 1.03 and 3.01 cd/m^2,

TABLE 5

Midpoints in cd/m^2 estimated from the pooled frequency data. Midpoints were obtained from a linear least squares analysis between the transformed frequencies and the variable stimulus values with ascending and descending frequencies pooled.

Subject	DM	JE	KM	LM1	LM2	LMP
Endpoints						
(a o b)	5.53	5.18	2.72	5.47	5.51	5.40
(c o d)	76.43	76.40	42.98	82.81	83.74	83.26
(a o c)	15.70	12.55	5.24	16.80	15.64	16.27
(b o d)	40.81	45.94	22.95	52.48	53.07	52.70
(a o b) o (c o d)	29.27	26.67	1.03	35.81		31.24
(a o c) o (b o d)	27.82	26.00	3.01	34.29		30.00

respectively. Both of these bisection points fall outside of the interval bisected, indicating that this subject had been unable to effectively follow the instructions and perform the task. Consequently, further consideration of this subject's performance will be confined to Appendix B and the present discussion will be limited to the results from the three remaining subjects.

The estimated bisection points obtained from the analysis of the frequency data for the remaining subjects are very similar, with all of LM's bisection points being slightly higher than those for the other two subjects.

Table 6 presents the estimated bisection points obtained when the frequencies from ascending and descending trials are analyzed separately. Here the estimated bisection points are more variable than when the frequencies were pooled, but they are once again very similar to each other. Notice that the estimated bisection points for ascending trials are considerably lower than those for descending trials. This is the reverse of the findings of Cross (1965) and Garner (1954) for loudness and Stevens (1957) for loudness and brightness. These investigators found that their subjects tended to adjust the bisection point on ascending trials slightly higher than on descending trials.

For purposes of comparing the present results with those from previous investigations, the estimated bisection points were converted into decibels (re: 10^{-6} cd/m^{2}).

TABLE 6

Midpoints in cd/m^2 estimated from the unpooled frequency data. Midpoints were obtained from a linear least squares analysis between the transformed frequencies and the variable stimulus values with ascending and descending trials analyzed separately.

Subjects	DM	JE	LM1	LM2	LMP
Endpoints					
(a o b) Ascending	4.38	5.79	4.75	4.81	4.76
Descending	7.28	5.80	5.72	5.82	5.98
(c o d) Ascending	71.80	73.15	75.71	76.40	76.12
Descending	83.96	80.39	92.64	100.24	95.77
(a o c) Ascending	13.99	10.91	12.84	13.83	13.49
Descending	19.01	14.12	21.27	17.33	19.89
(b o d) Ascending	35.54	38.35	44.61	46.31	45.94
Descending	47.58	53.46	60.36	59.04	59.70
(a o b) o (c o d)					
Ascending	26.44	20.61	28.55		24.95
Descending	32.53	32.79	39.95		35.62
(a o c) o (b o d)					
Ascending	26.34	21.91	29.38		26.25
Descending	29.91	29.35	46.84		33.97

These are presented in Table 7. Subtracting the estimated bisection points for ascending trials from those for descending trials gives differences that reflect the differential temporal order effects of the two types of trials. Table 8 presents these differences. The differences between the bisection points for ascending versus descending trials are very small: the mean is 1.19 decibels and the standard deviation is .51 decibels.

Stevens (1957) reported that on ascending trials his subjects tended to overestimate the bisection point of a loudness interval by 5-8 decibels with a somewhat smaller overestimation for brightness. He suggested that the differences for brightness are smaller than those for loudness because the range of stimuli was smaller: 50-80 decibels in the brightness experiment compared with 60-100 decibels in the loudness experiment. Because the decibel scale is logarithmic, the 20 decibel range from 61-81 decibels in the present research is smaller than the ranges in Stevens' experiments, but it is not clear that a smaller range is subjectively larger or smaller than another. A 20 decibel range from 60-80 decibels may be subjectively quite different than a 20 decibel range from 80-100 decibels.

It would appear that using the method of constant stimuli greatly reduces the amount of the difference between bisection points for ascending versus descending trials over the method of adjustment. However, even though the differences were small, their presence suggests that

TABLE 7

The estimated bisection points from Tables 5 and 6 transformed into decibels (re: 10^{-6} cd/m^2).

Subject		DM	JE	LM1	LM2	LMP
Endpoints						
(a o b)	Ascending	66.42	67.63	66.77	66.82	66.78
	Descending	68.62	67.63	67.57	67.65	67.77
	Pooled	67.42	67.14	67.38	67.41	67.32
(c o d)	Ascending	78.56	78.64	78.79	78.83	78.82
	Descending	79.24	79.05	79.67	80.01	79.81
	Pooled	78.83	78.83	79.18	79.23	79.20
(a o c)	Ascending	71.46	70.38	71.09	71.41	71.30
	Descending	72.79	71.50	73.28	72.39	72.99
	Pooled	71.96	70.99	72.25	71.94	72.11
(b o d)	Ascending	75.51	75.84	76.49	76.66	76.62
	Descending	76.77	77.28	77.81	77.71	77.76
	Pooled	76.11	76.62	77.20	77.25	77.22
(a o b) o (c o d)						
	Ascending	74.22	73.14	74.56		73.97
	Descending	75.13	75.16	76.02		75.52
	Pooled	74.66	74.26	75.54		74.95
(a o c) o (b o d)						
	Ascending	74.21	73.41	74.68		74.19
	Descending	74.76	74.68	76.61		75.31
	Pooled	74.44	74.15	75.35		74.77

TABLE 8

Differences between ascending and descending bisection points from Table 7 (descending minus ascending) in decibels.

Subject	DM	JE	LM1	LM2	LMP
Endpoints					
(a o b)	2.20	0.00	0.80	0.83	0.99
(b o d)	0.68	0.41	0.88	1.18	0.99
(a o c)	1.33	1.12	2.19	0.98	1.69
(b o d)	1.27	1.44	4.24	1.05	1.14
(a o b) o (c o d)	0.91	2.02	1.46		1.55
(a o c) o (b o d)	0.55	1.27	1.93		1.12

commutativity was violated in the bisection task using the method of constant stimuli. The direction of this difference is opposite from the results of previous research. Perhaps this reversal was due to the manner in which the standard and variable stimuli were presented in the present research. The variable stimulus was presented after the two standards instead of between them thus eliminating the staircase effect of presenting the variable stimulus between the endpoints. Therefore, it seems possible that the usual finding that the ascending bisection point is adjusted higher than the descending bisection point may be due to this staircase effect or to the method of adjustment itself.

The Psychophysical Function

Once the bisection points have been estimated, the psychophysical function can be investigated. Since the stimuli were presented sequentially, and since commutativity does not seem to hold, an appropriate expression would weight the endpoints differentially depending upon whether the trial was in ascending or descending order. The expression fitted to the data was:

$$X_i = \left[wX_h^k + (1 - w)X_j^k \right]^{1/k} , \qquad (9)$$

where X_i is the estimated bisection point, X_h and X_j refer to the first and second presented endpoints instead of the lowest and highest, and w is a weighting associated with

the temporal effects of the sequential presentation of the stimuli. Thus, w is associated with the first presented endpoint while (1 - w) is associated with the second presented endpoint, regardless of whether it was an ascending or descending trial. In fitting this expression it was assumed that the bisection point is an average of the two standards, but the standards are weighted differentially depending upon the temporal order of their presentation. It was also assumed that reflexivity was not violated, and hence that the weights sum to unity.

Equation 9 was fitted to the data by the Gauss-Newton method (Hartley, 1964), an iterative least squares procedure for solving nonlinear equations. Two sets of estimates of the paramters w and k were obtained. One set was obtained by minimizing the sum of the unweighted squared deviations of the obtained estimated bisection points from those predicted by the solution, and the other set was obtained when these deviations were weighted by $1/X_i^2$. This weighting minimizes the sum of the squares of the relative deviations of the observed from the predicted. The weighted solution was used because it is often found that variability increases as stimulus magnitude increases.

The expression was fitted to the bisection points estimated from ascending trials, descending trials, and simultaneously to both sets combined but not to the data from ascending and descending trials pooled (i.e., Table 5). An analysis of the pooled data would be of dubious value

because the parameter w would not reflect the effects of the sequential presentation of the stimuli as in the other analyses. Solutions were obtained from the data from JE, DM and LMP and from these sets of data combined in a single analysis. The parameter estimates for the weighted and unweighted solutions are presented in Table 9.

The estimates of k from the solutions obtained simultaneously for ascending and descending trials gave estimates of the sensory or input parameter of .409 and .327 over the three subjects for the weighted and unweighted solutions respectively. The first estimate is somewhat larger than the usual value of .33 obtained from magnitude estimates of brightness, while the second is in close agreement with that value. However, it is about twice the value of k obtained from interval judgment procedures (Curtis, 1970) and it is larger than the k values that Fagot & Stewart (1970) obtained, although it is difficult to interpret many of their estimates since some were negative.

The weighted and unweighted estimates of k for the data from individual subjects for ascending and descending trials separately exhibit some variability, but there are no consistent trends. When the analysis was performed on the data for all three subjects combined, k is larger for ascending trials than for descending trials: .57 and .497 for ascending and .368 and .289 for descending for the weighted and unweighted solutions respectively.

TABLE 9

Parameter estimates from a nonlinear least squares curve fitting solution of Equation 9 to the frequency data. A weighted and unweighted solution was obtained for each subject and for subjects combined. The analysis was done separately for the data from ascending and descending trials, and for these two sets of data combined.

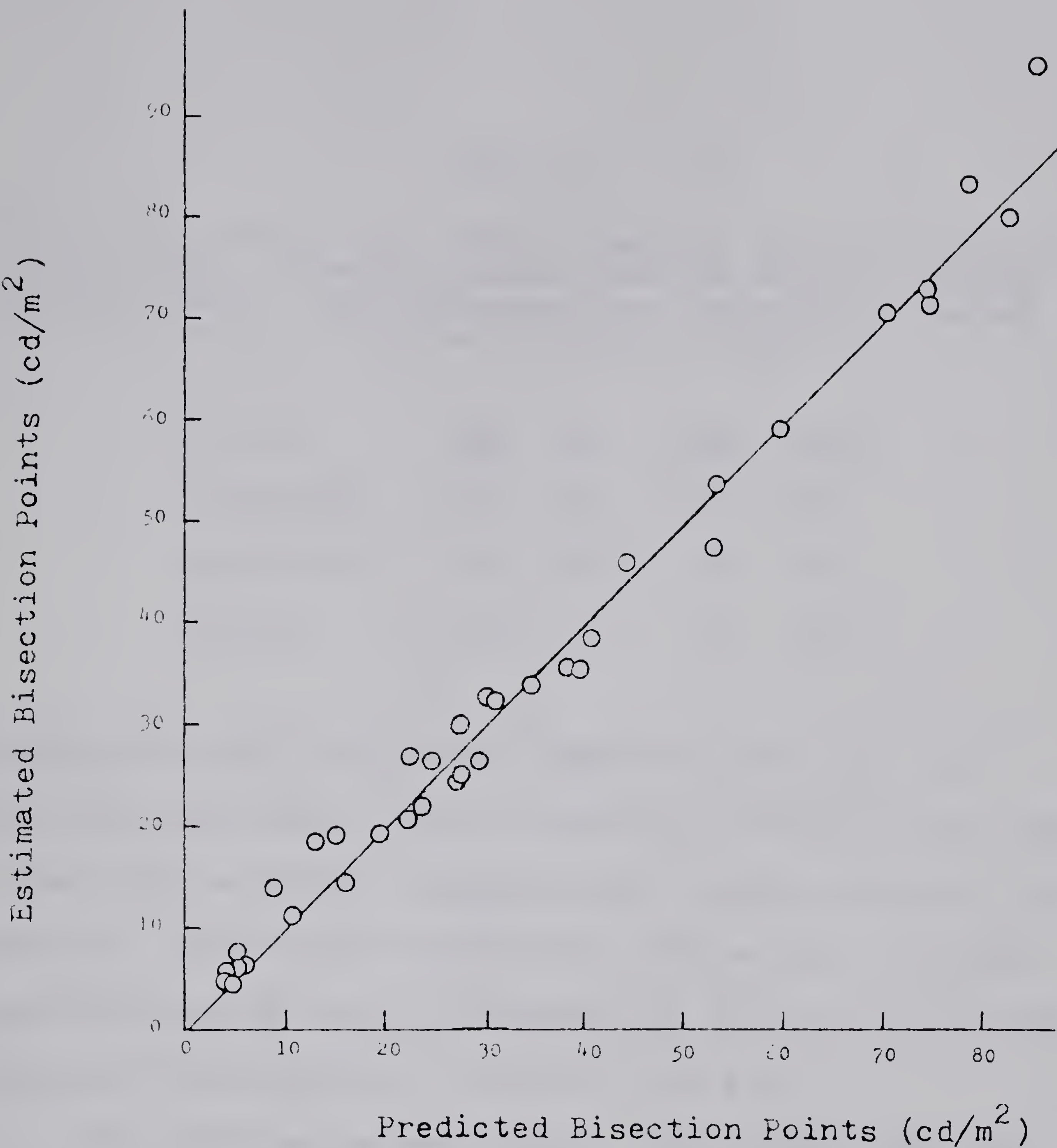
Parameters	Ascending Trials		Descending Trials		Combined	
	w	k	w	k	w	k
Subjects						
Weighted Solution						
DM	.665	.768	.545	.491	.564	.443
JE	.588	.413	.640	.876	.560	.325
LMP	.615	.635	.599	.398	.571	.495
All 3	.616	.570	.607	.368	.563	.409
Unweighted Solution						
DM	.643	.595	.544	.352	.563	.296
JE	.598	.381	.525	.453	.571	.292
LMP	.574	.502	.708	-.005	.581	.493
All 3	.605	.497	.598	.289	.559	.327

The effects of sequentially presenting the endpoints is apparent. The weights are all greater than .5 indicating that the first endpoint presented was given more weight than the second, independent of whether it was an ascending or descending trial. Furthermore, in four of the six cases, the weight for ascending trials is greater than for descending trial.

The predicted bisection points from the fit of Equation 9 to the data of all three subject with ascending and descending trials combined ($n = 36$, $w = .559$, $k = .327$) are graphed in Figure 10 against the bisection points estimated from the least squares analysis of the frequency data (Table 6). There are systematic deviations between the estimated bisection points and those predicted by Equation 9. The function tends to underestimate the bisection point for lower intensities and overestimate for higher intensities. This may be due to inaccuracies in the form of the psychophysical function assumed to describe the data, or it may be due to violations of reflexivity.

The Bisymmetry Axiom

The estimated bisection points of import, those needed for a test of the bisymmetry axiom are given in the last six lines of Table 7. According to the bisymmetry axiom, these final bisection points should be the same within experimental error. The differences



$$X_i = .559X_h^{.327} + (1 - .559)X_j^{.327} \quad 1/.327$$

Figure 10: Estimated bisection points as a function of values predicted by a fit of Equation (1) to the estimated bisection points from the transformed frequency data.

TABLE 10

Differences between the final bisection points needed for a test of the bisymmetry axiom from the analysis of the frequency data. The differences are from the estimated bisection points in decibels given in Table 7.

Subject	DM	JE	LM1	LMP
Ascending	.01	-.27	-.12	-.22
Descending	.37	.48	-.59	.21
Pooled	.22	.11	.19	.18

between the final estimated bisection points for the data from ascending trials, descending trials and the data pooled over ascending and descending trials are given in Table 10. These were obtained by subtracting the final bisection points for the bisection (a o c) o (b o d) from those for the bisection (a o b) o (c o d).

The differences are slight. The overall mean and standard deviation of the absolute differences were .248 and .047 decibels respectively while the overall mean and standard deviation of the signed differences were .048 and .301 respectively. These differences are much smaller than the differences obtained by Garner (1954) or Gage (1934a) and are similar to the results obtained by Newman, et al. (1937) for loudness. They are also smaller than the differences obtained by Cross (1965) for loudness. Although the present study was concerned with brightness, the comparison can be justified because loudness and brightness

have similar input or sensory parameters in the psychophysical function and they have similar thresholds (cf., Stevens, 1955).

To verify that the values in Table 10 are extremely precise, an additional analysis was performed. The data from subject LMP was divided into two sets and the least squares analysis was performed on each set. The final bisection points from this analysis are given in Table 11.

TABLE 11

The final bisection points needed for a test of the bisymmetry axiom and their differences from a split half analysis of the data of LMP. The data have been pooled over ascending and descending trials for the two halves.

	First Half	Second Half	Difference
(a o c) o (b o d)	73.357	75.352	-1.995
(a o b) o (c o d)	75.567	73.914	1.653
Difference	2.285	-1.437	

This Table can be inspected in two different ways. First, the difference between the estimated bisection points for the two halves of the data can be viewed as an index of the reliability of the responses of a single subject. These difference of -1.995 and 1.653 are of a much larger magnitude than the differences in Table 10, but are similar in magnitude to the differences between ascending and descending trials given in Table 8.

Second, the differences between the final bisection points needed for a test of the bisymmetry axiom from the two halves of the data can be viewed as an index of the reliability of the estimates of the two crucial bisection points. For the first half, these differ by 2.285, whereas for the second half they differ by -1.437 in the opposite direction. Those within subject differences are many times larger than those listed in Table 10, suggesting that the latter are about as accurate as could be expected, considering error.

The method of constant stimuli employed in the present experiment appears to have eliminated some of the sources of bias present in the method of adjustment: those due to the initial placement of the adjustable stimulus and other less definite sources. The bias due to ascending or descending orders of presentation of the endpoints is present in the data obtained here, but that bias did not preclude the successful empirical verification of the bisymmetry axiom.

Reaction Time Data

The latency of the choice response for each variable stimulus associated with the various endpoint pairs were averaged by taking harmonic means. These are graphed in Figures 11-15. Harmonic mean reaction time pooled over ascending and descending trials and harmonic mean reaction time for ascending and descending trials separately are presented for each pair of endpoint stimuli for each subject.

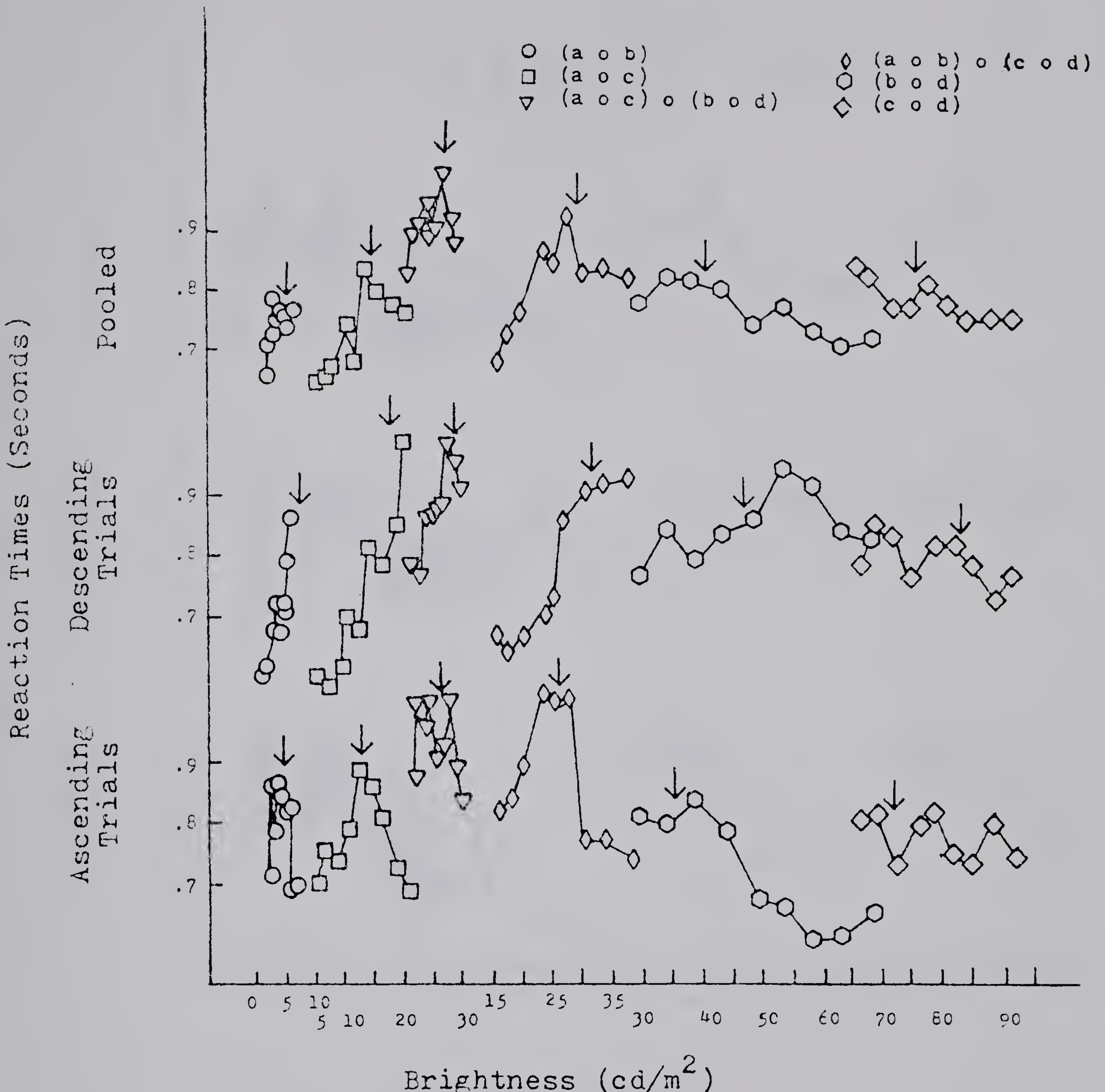


Figure 11: Reaction times as a function of brightness for subject DM. Lines connect reaction time for the variable stimuli associated with each pair of endpoints. Arrows indicate the estimated bisection points obtained from the analysis of the frequency data.

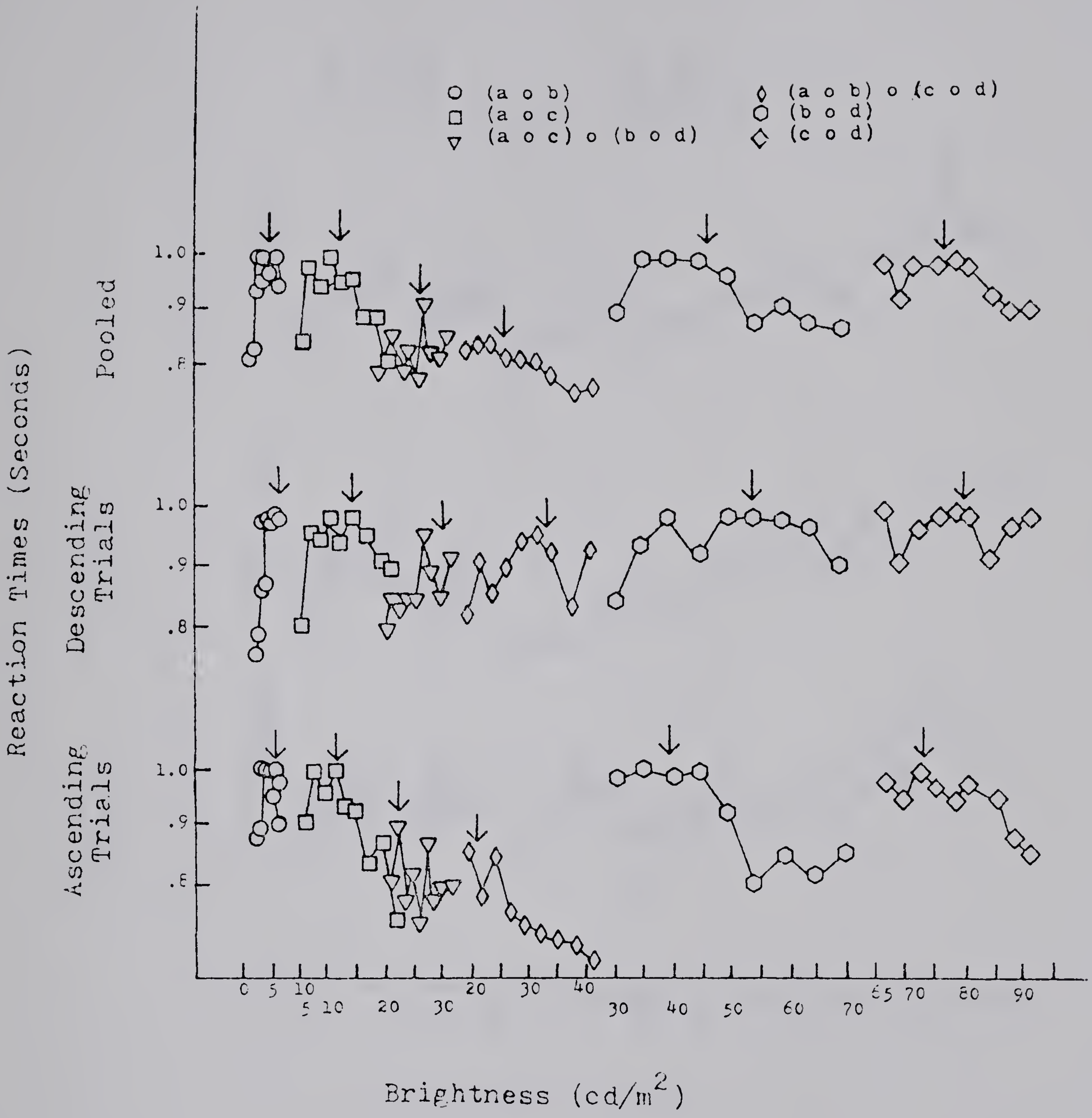


Figure 12: Reaction times as a function of brightness for subject JE. Lines connect reaction times for the variable stimuli associated with each pair of endpoints. Arrows indicate the estimated bisection points obtained from the analysis of the frequency data.

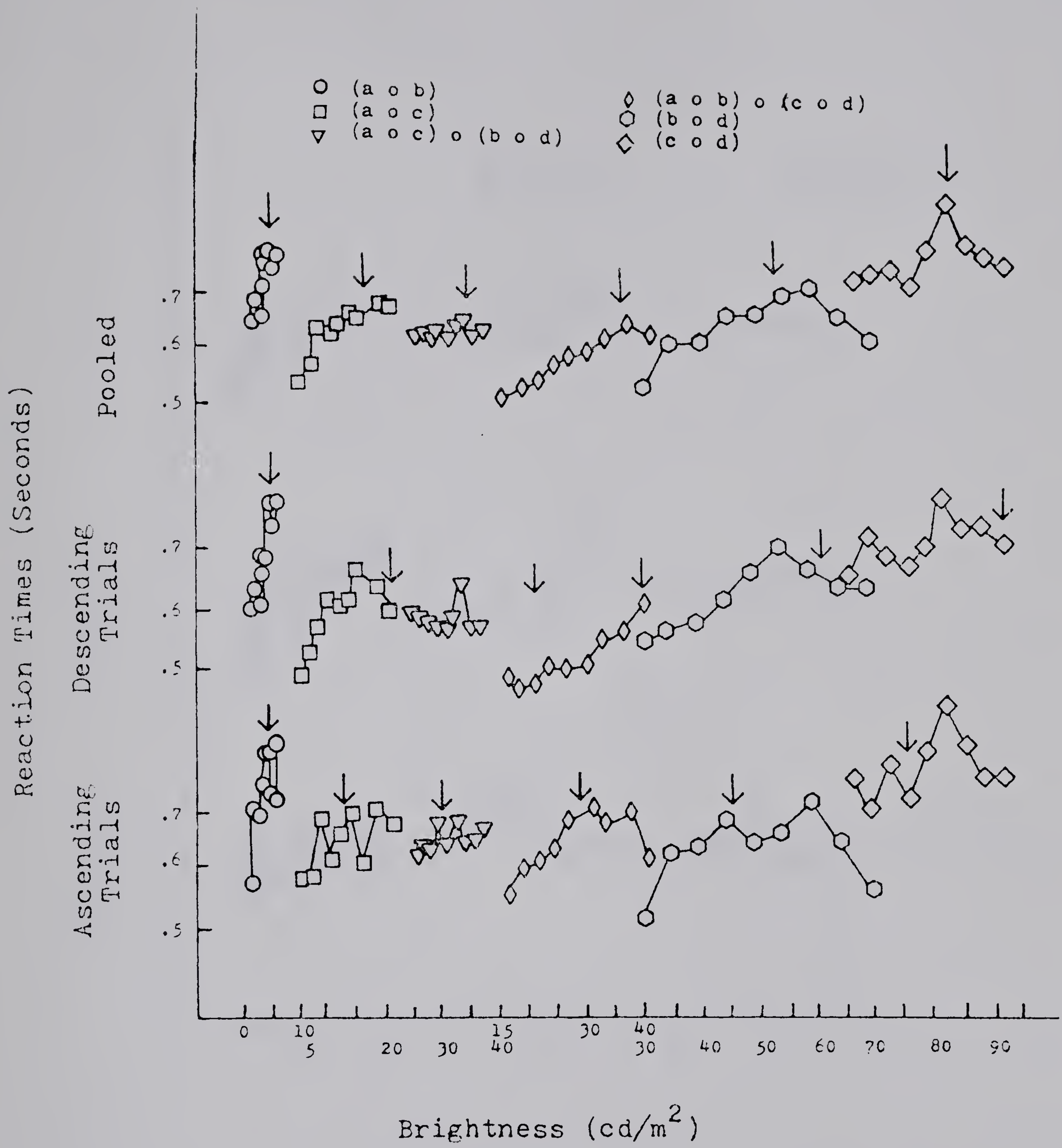


Figure 13: Reaction times as a function of brightness for subjec LM1 (i.e., the basic experiment). Lines connect reaction times for the variable stimuli associated with each pair of endpoints. Arrows indicate the estimated bisection points obtained from the analysis of the corresponding frequency data.

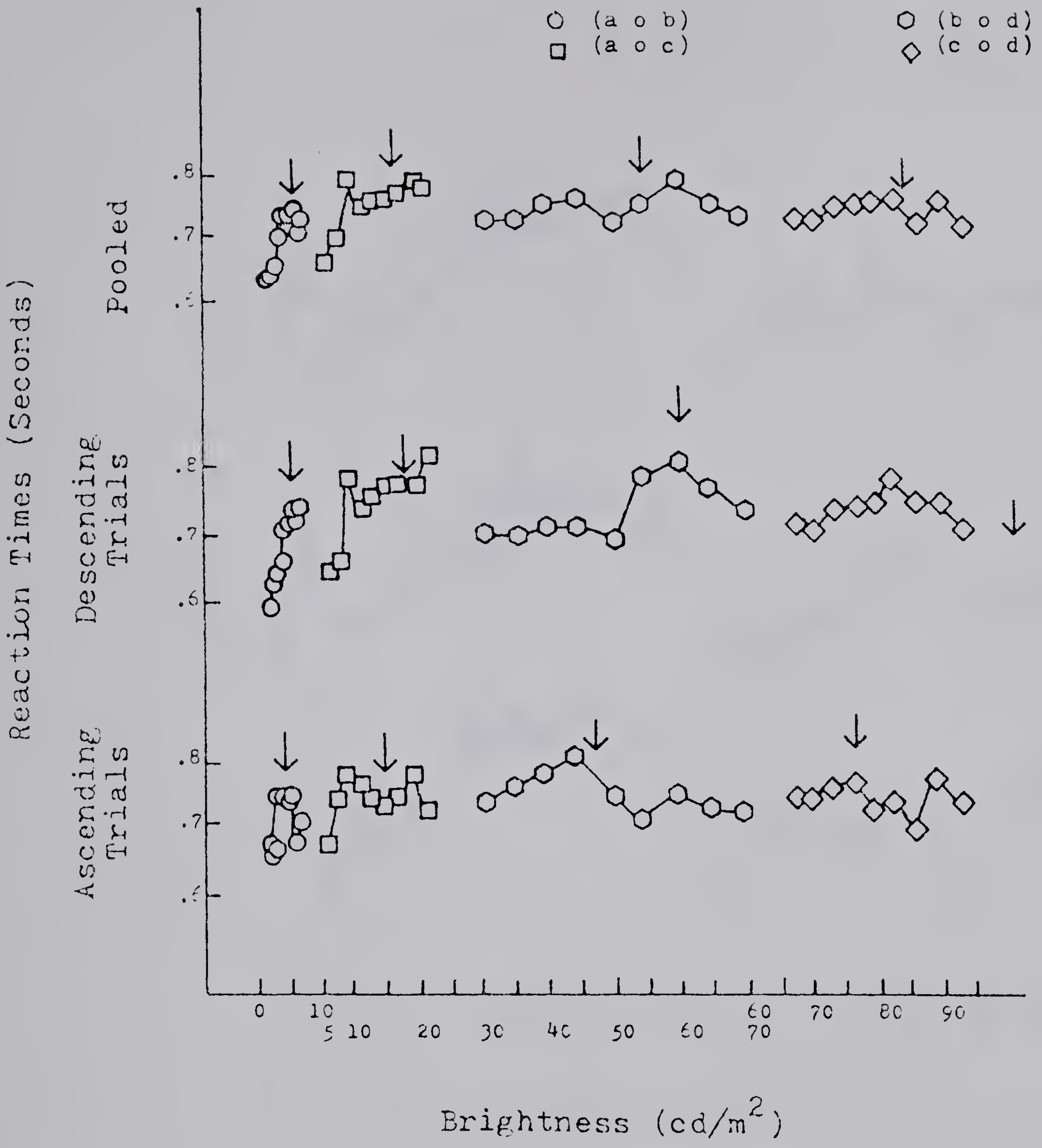


Figure 14: Reaction times as a function of brightness for subject LM2 (i.e., the first part of the replication). Lines connect reaction times for the variable stimuli associated with each pair of endpoints. Arrows indicate the estimated bisection points obtained from the analysis of the corresponding frequency data.

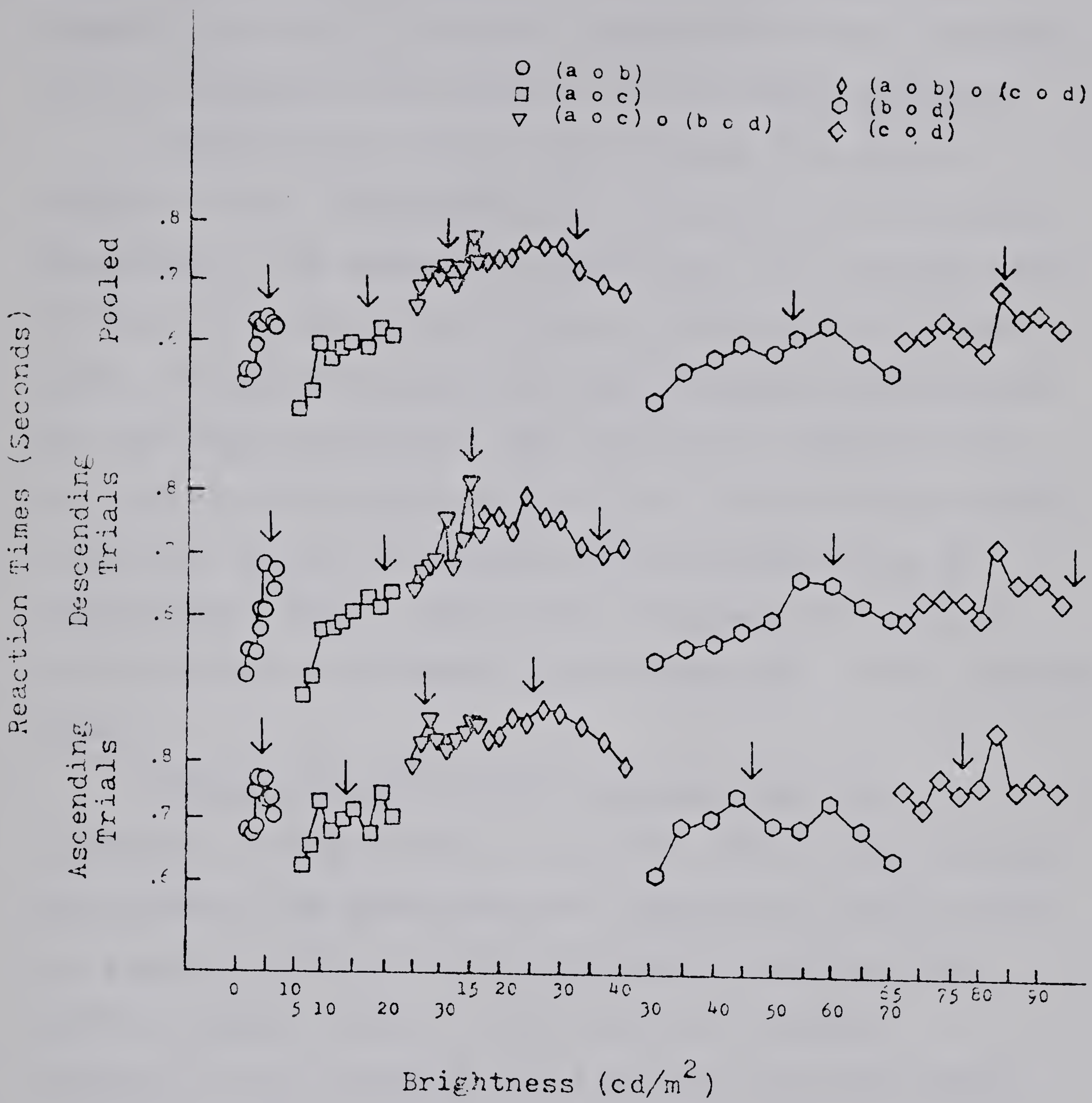


Figure 15: Reaction times as a function of brightness for subject LMP. The plots represent the data of Figure 14 combined with the comparable data of Figure 13, along with the data from the last part of the replication (the final bisections). Lines connect reaction times for the variable stimuli associated with each pair of endpoints. Arrows indicate the estimated bisection points obtained from the analysis of the corresponding frequency data.

The different symbols represent the harmonic mean reaction time for the variable stimuli associated with the different pairs of endpoints as indicated in the figure captions.

Figures 13-15 present the reaction time data for subject LM who participated in a replication of the basic experiment. The format for presenting this subject's data here is the same as for the presentation of his frequency data. Figure 13 presents the harmonic mean reaction time for the basic experiment (LM1), Figure 14 presents them for the second 25 replications of the first four bisections (LM2), and Figure 15 presents the pooled data from 50 replications (LMP). The arrows represent the bisection points that were estimated from the analysis of the frequency data.

Subjects tended to have the same range of latencies -- from 0.6 to 1.0 seconds. This range is somewhat higher than choice reaction times from judging which of a pair of lights is brighter (Mullin, Curtis & Rule, 1978). Mullin, et al. (1978) obtained reaction times of .4 to .6 seconds for both white and red lights. In addition, other research has shown that the wavelength of the stimuli do not affect reaction times to judgments about intensity if there are intensity cues available on which to base a judgment (cf., Nissen & Pokorny, 1977). Instead, the longer reaction times in the present research would seem to be attributable to the increased difficulty of the present task. Subjects were making judgments

about the difference of two intervals instead of the difference of two stimuli.

The graphs are generally of an inverted U-shaped form. It is apparent that longer reaction times are associated with variable stimuli that are close to the bisection points that were estimated from the frequency data (indicated by arrows), and choice reaction time decreased as the difference between the variable stimulus and the bisection point increased. This is consistent with the interpretation that choice reaction time is inversely related to the subjective difference between the bisection point and the variable stimulus.

The Psychophysical Function

If the assumption that a choice reaction time is inversely related to the difference between the subjective magnitude of the two intervals is correct, Equation 7 may provide a description of the relation between reaction time and stimulus values.

Equation 7 was fitted to the group data by the Gauss-Newton method in two different analyses. In one analysis, the dependent variable consisted of the set of harmonic mean reaction times over both ascending and descending trials for each of the three subjects (Pooled). In the second analysis, the reaction times from ascending trials were not pooled with those from descending trials, but were considered simultaneously with them in the least

squares solution (Combined). The best fitting equations from the two analyses are given by:

$$\text{Pooled} = -.132(\, X_h^{.271} + X_j^{.271} - 2X_i^{.271}\,)^{.71} + .869\,,$$

and

$$\text{Combined} = -.127(\, X_h^{.275} + X_j^{.275} - 2X_i^{.275}\,)^{.74} + .872\,,$$

where X_h and X_j are the endpoint stimuli and X_i is the variable stimulus.

Several features of these equations may be noted. First, the value of the sensory or input exponent k appears to be in the neighborhood of .3 and is similar in magnitude to the exponent for brightness postulated by S. S. Stevens and his associates on the basis of results from magnitude estimation and other ratio judgment procedures. Second, neither empirical equation is consistent with the assumption that choice reaction time is reciprocally related to the difference between subjective intervals according to Equation 7. If the reciprocal relation had held, the estimate of the exponent m in Equation 7 should have been negative and close to unity, and the coefficient a would have been positive. Neither is true for the present data. In addition, the constant b in the above expressions are too large in value to be interpretable as an irreducible minimum reaction time, beyond which no further reduction can be achieved. Considering that the actual data ranged from .6 to 1.0 seconds, the estimated values of .869 and .872 seconds do not suggest interpretation as an

irreducible minimum.

A graph of the reaction times predicted by the expression against the obtained reaction times are given in Figures 16 and 17 for the two solutions. Inspection reveals that the predicted reaction times are all between .7 and .9 seconds. Clearly, the obtained solutions do not describe the data closely, nor do the parameters have any meaningful psychological interpretation.

There is an obvious explanation. Notice that the graphs of the harmonic mean reaction times for ascending and descending trials often peak at different stimulus values (Figures 11-15), suggesting a violation of commutativity similar to the one found with the frequency data. Consequently, Equation 7 is not the appropriate model to describe the data. The model should have weights associated with each of the independent variables and these weights should sum to unity if reflexivity is not violated. However, attempts to obtain a solution to the more appropriate model were not successful. This may have been due to complex interactions in the data such that the model that would describe the data is more complicated than a simple weighted model. Other research involving reaction times to pairs of sequentially presented visual stimuli has suggested that there are interactions that are not accounted for by a simple weighted average of the pair of stimulus intensities (Mullin, Curtis & Rule, 1978).

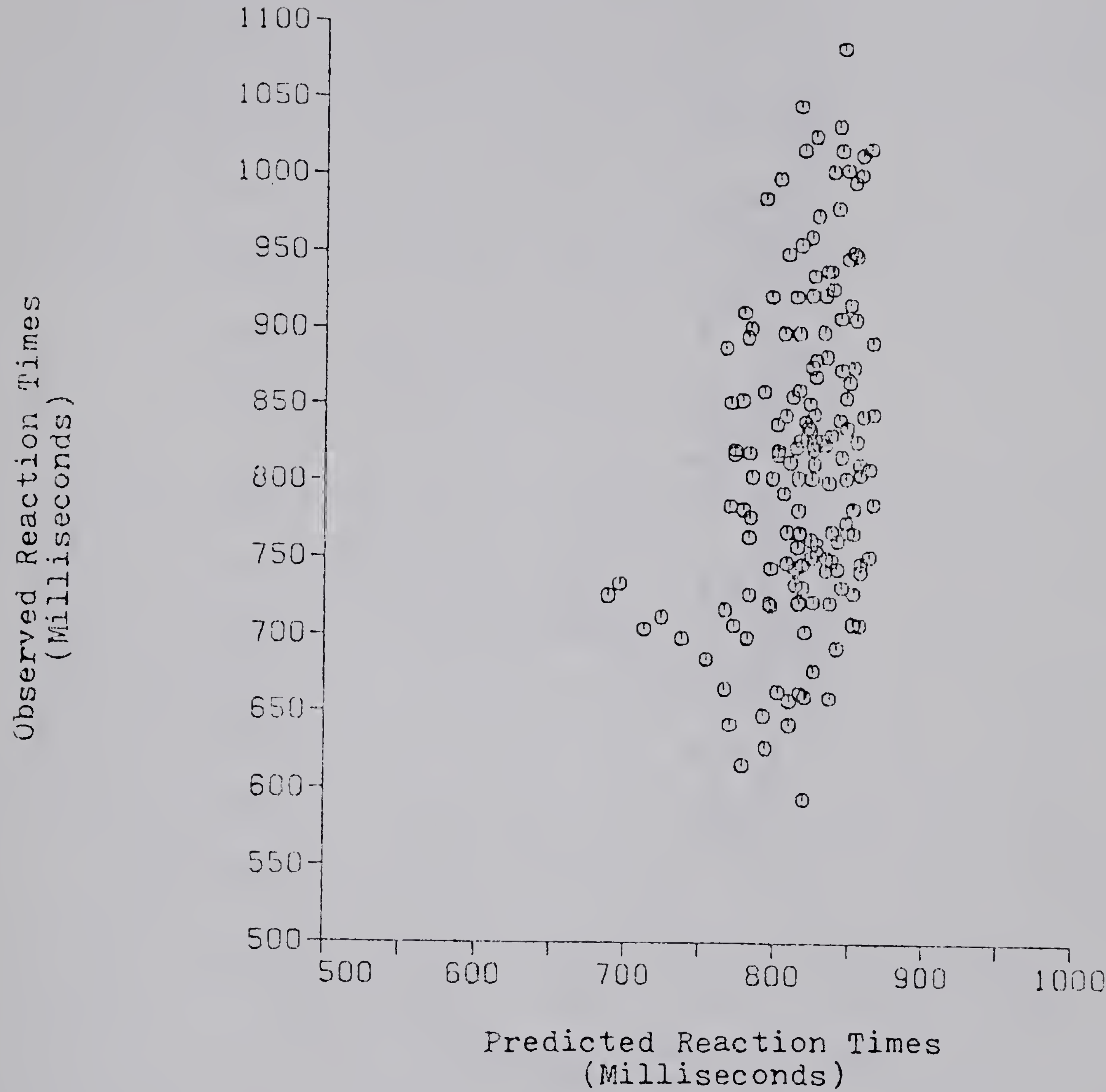


Figure 16: Harmonic mean reaction times as a function of values predicted by a fit of Equation (7) to the reaction times pooled over ascending and descending trials (n = 162).

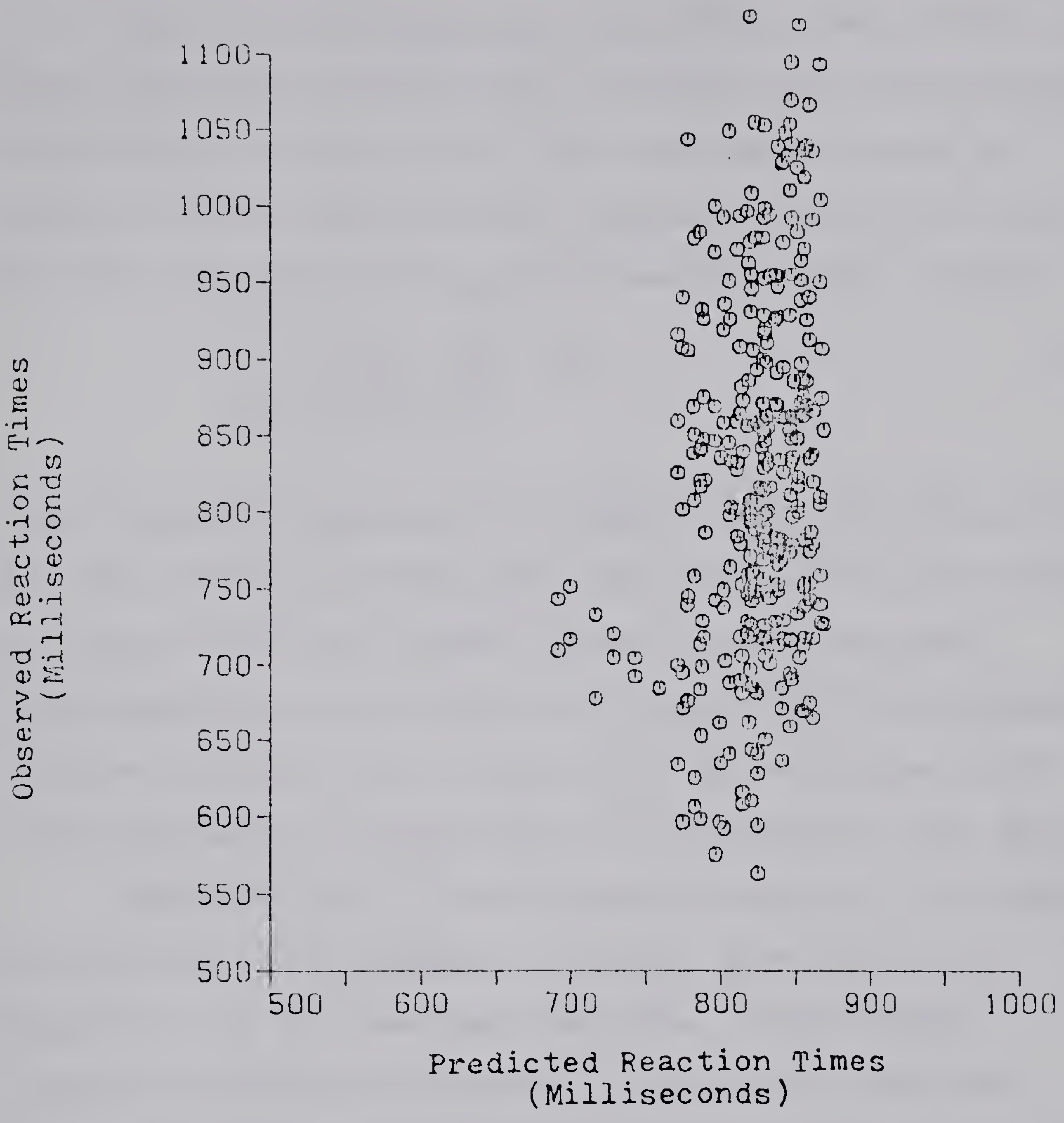


Figure 17: Harmonic mean reaction times as a function of values predicted by a fit of Equation (7) to the reaction times for ascending and descending trials combines (n = 324).

The Bisymmetry Axiom

Once the psychophysical function has been fitted to the data, bisection points can be estimated from the reaction time data using Equation 7. The expression inside the absolute values signs in that expression can be set equal to zero and solved for X_i, the bisection point, giving:

$$X_i = \left[\frac{X_a^k + X_b^k}{2}\right]^{1/k} \quad . \tag{10}$$

This expression assumes that commutativity and reflexivity are not violated and that the bisection point is the average of the two endpoint stimuli raised to the kth power. Unfortunately, commutativity is violated and the parameters obtained from the fit of Equation 7 do not suggest that this expression is appropriate to the reaction time data.

Nevertheless, it was decided to estimate the bisection points using the parameters obtained from the fit of Equation 7 to the reaction time data. Substituting endpoint stimuli into Equation 10 gives the bisection points for the various bisections required for a test of the bisymmetry axiom. These are presented in Table 12. The estimated bisection points for the initial four bisections are the same for all subjects because the estimates of parameters of Equation 7 were obtained from group data and the endpoints for these bisections were the same for all subjects. The bisection points were estimated using the parameters of both of the fits of Equation 7 to the data,

TABLE 12

Estimated bisection points in decibels from Equation 10 using the parameters estimated from a fit of Equation 7 to the harmonic means of the reaction times. The analysis was done for pooled reaction times and for the reaction times from ascending and descending trials combined.

	Combined	Pooled
Endpoints		
(a o b)	66.46	66.45
(b o c)	78.95	78.95
(a o c)	70.76	70.74
(b o d)	76.54	76.53
(a o b) o (c o d)		
DM	74.14	74.12
JE	74.047	74.03
LMP	74.30	74.34
(a o c) o (b o d)		
DM	74.17	74.17
JE	74.05	74.05
LMP	74.87	74.87

that is, for the fit of the function to the pooled reaction times and simultaneously to the reaction times from ascending and descending trials combined.

An inspection of Table 12 in comparison to the pooled bisection points from the frequency data in Table 7, reveals a striking correspondence: Even though the fit of Equation 7 to the reaction time data was not at all acceptable, the bisection points estimated from the parameters obtained converge quite nicely with the estimated bisection points obtained from an analysis of the frequency data.

The estimated bisection points of import, those needed for a test of the bisymmetry axiom are given in the last 6 lines of Table 12 for individual subjects. According to the bisymmetry axiom, these final estimated bisection points should be the same within experimental error for each subject. The differences between these final estimated bisection points for the two different psychophysical function analyses (Pooled and Combined) are given in Table 13. The differences are minimal. The overall mean and standard deviation of the absolute differences are .199 and .111 decibels respectively while the overall mean and standard deviation of the signed differences are -.198 and .111 decibels respectively. These differences are similar to those obtained from the analysis of the frequency data and also to those of Newman, et al. (1937). However, they are smaller than the findings of Garner (1954) and Gage (1934a). Again, while the fit of the function of Equation 7 was not

TABLE 13

Differences in decibels between the final estimated bisection points needed for a test of the bisymmetry axiom using the reaction time data. The differences are given for the Combined and Pooled analysis presented in Table 12.

Subject	DM	JE	LMP
Combined	-.031	.003	-.570
Pooled	-.043	-.018	-.528

very good, it gives estimated bisection points that are consistent with the support for the bisymmetry axiom that was obtained from the analysis of the frequency data.

Because the model of Equation 7 did not describe the data separately for ascending and descending trials, a quantification of the bias introduced by these two types of trials could not be obtained as it was with the frequency data. Assuming, however, that it is present, it did not affect the verification of the bisymmetry axiom using the bisection points estimated from the psychophysical function.

The method of constant stimuli applied to the bisection experiment has demonstrated value over the method of adjustment. The frequency data support the bisymmetry axiom, independent of whether temporal order or sequential effects are present. The paradoxical point, however, is the fact that the analysis of the reaction time data also supports

the bisymmetry axiom in spite of the inadequacies of the psychophysical function that was assumed to underly those data.

DISCUSSION

The results of the analysis of the data lend some empirical support to the bisymmetry axiom and to the bisection model in measurement. This was particularly the case with the results of the analysis of the frequency data. Although some results from the analysis of the reaction time data support the bisymmetry axiom and converge with the bisection points obtained from the analysis of the frequency data, the findings are based on an expression where the fit to the data is poor.

In the analysis of the frequency data, it was assumed that the relationship between variable stimulus values and the transformed frequencies was linear and the bisection point was assumed to be that stimulus value which was judged to be more similar to the brightest endpoint 50% of the time. The differences between the two final bisection points required for a test of the bisymmetry axiom were minimal and their absolute values averaged .248 decibels. These differences are smaller than the differences obtained by previous researchers using the method of adjustment (Garner, 1954; Gage, 1934a). It would appear that the method of constant stimuli eliminates some sources of error inherent in the method of adjustment such as the initial placement of the adjusted stimulus.

Newman, et al. (1937) found that the bisection points of intervals that were large with respect to the sone scale of loudness tended to be underestimated relative to the bisection points of intervals that were smaller in terms of the sone scale. A similar result was evident for bisections of brightness intervals obtained by Stewart, Fagot and Eskildsen (1967). Their results imply that if the psychophysical relation may be described by a power function, the magnitude of the exponent decreases as the relative interval size increases. In terms of the present research, that implies that the bisection represented by (a o b) o (c o d) should yield a lower midpoint than the bisection indicated by (a o c) o (b o d). The opposite appears to be the case for the frequency data pooled over ascending and descending trials; however, when the bisection points yielded by ascending and descending trials are considered separately, these differences do not show a systematic bias (see Table 10). Thus, it is not entirely clear what effect the size of the interval has on the bisection point.

The differences between bisection points obtained from ascending and descending trials that is usually found in the bisection task are evident in the present results. Even though the differences are small, ascending midpoints were consistently estimated to be below descending midpoints. This indicates a violation of commutativity, although violations of reflexivity also may be involved. However,

Pfanzagl (1968) and Fagot & Stewart (1970) found that in the bisection task commutativity is usually violated while reflexivity is not.

In any event, the finding of an underestimation of the midpoint for ascending trials is opposite to the findings of Stevens (1957). He found that his subjects tended to set the midpoint 5-8 decibels higher on ascending trials for loudness, while for brightness the overestimation was reported to be somewhat smaller. Cross (1965) also found that subjects tend to set the midpoints somewhat higher on ascending trials in a direct test of the bisymmetry axiom using the method of adjustment. There are several differences between the experiments of Stevens and Cross and the present research which may account for the reversal of the bisection points for ascending and descending trials.

First, Cross presented all of the ascending trials in the first half of the experiment and all of the descending trials in the second half. His experiment used sequentially presented auditory stimuli and the method of adjustment. The present research used the method of constant stimuli for brightness bisection, and the two types of trials were randomized within blocks. The different methodologies and procedures for randomization and the different sensory modalities could have produced differences in how the subjects perceived the two types of trials. Stevens (1957)

reported that he used the method of adjustment in his bisection experiment, but the details of procedure and method of randomization were not reported. His finding of less hysteresis for brightness than for loudness was attributed to the shorter range used in the brightness study. However, Cross (1965) used a shorter range than Stevens in his loudness study, but he obtained differences between ascending and descending trials that were similar to Stevens' differences (4.2 decibels). The source of this bias is not altogether clear or simple. When auditory stimuli are sequentially presented, stimuli presented early in the sequence probably do not affect the preception of sebsequent stimuli greatly (Marks, 1974), but this is clearly not the case with sequentially presented visual stimuli for which there may be pronounced effects of adaptation or masking. Yet, the bias due to ascending and descending trials is greater for loudness than brightness.

Another factor may be that in most previous research the variable stimuli were presented between the standards creating an ascending or descending staircase effect. In the present research the variable stimuli were presented after the two standards instead of between them. This methodological difference may have reversed the effects of ascending and descending presentations.

The results of a study by Mullin, Curtis and Rule (1978) suggest one explanation for the present finding that the bisection point for descending trials is judged

as greater than the bisection point for ascending trials. Mullin, et al. measured the time required by subjects to judge which of a pair of sequentially presented visual stimuli was the brighter by pressing one of two buttons. In their incomplete factorial design, there were ten pairs of stimuli which were presented in both ascending and descending order. Although not statistically significant, the harmonic mean reaction time for the descending trials was .532 seconds and for the ascending trials it was .522 seconds.

It has frequently been suggested that choice reaction time is inversely related to the perceived difference between stimuli. If such is the case, these results suggest that the two stimuli are perceived as more similar when they are presented in descending order than when they are presented in ascending order. Suppose that a given stimulus is perceived as brighter when it is presented second than when it is presented first in the sequence. Such a result might be a consequence of a fading trace of the first stimulus held in memory or backward masking, among other possible explanations. The consequence would be that the less intense second stimulus in a descending trial should be enhanced resulting in a relatively small perceived difference, whereas the increase in apparent intensity of the brighter second stimulus in ascending trials should increase the apparent separation of the two stimuli. In fact, the harmonic mean reaction time for all descending

trials in the Mullin, et al. study was .522 seconds, while the comparable mean for ascending trials was .513 seconds. (Although no statistical test was performed, these means are likely not significantly different.)

The foregoing result is consistent with the hypothesis that the perceived difference between two stimuli may differ depending on the order in which they are presented; however, an alternative interpretation can also be suggested. That is, the difference between the reaction times for ascending and descending trials may reflect only the processing of information concerning the magnitude of the second stimulus, as in a simple reaction time task. It is well known that the time in responding to a visual or auditory stimulus varies inversely with the physical measure of the stimulus (Stevens, 1975; McGill, 1963). Presumably, latency of response reflects the time required for the processing of intensity information. Consequently, the faster response in the case of ascending trials may reflect nothing more than the fact that the time required for a subject to respond to a bright second stimulus is shorter than is required for him to respond to one that is less bright (assuming, of course, that the total time for a trial is more than sufficient so that the processing of information concerning the first stimulus is not a significant factor).

The Mullin, et al. analysis was based only on the data from correct trials. When an error occurred, additional trials were included so that an equal number of correct observa-

tion were available for all stimulus pairs. However, an inspection of the error data reveals some interesting trends. There were more errors on ascending trials than on descending trials and the number of errors increased as the intensity of the first stimulus increased. Thus, for any one stimulus pair, on ascending trials the second stimulus is seen as dimmer than on descending trials and the subject makes more errors in judgment. In addition, for stimulus pairs in which the stimuli were identical, reaction time decreased as the absolute intensity of the pair increased. The number of times the second stimulus of the pair was judged as brighter decreased from approximately 50% to about 33% as the absolute intensity of the pair increased. Thus, for sequentially presented visual stimuli, the responses reflect an interaction of several factors. The reaction time data suggest that the order in which the stimuli are presented affects the response, and the error data support this, since there were more errors for ascending trials than descending trials. The sequential effects interact with the effects of the absolute intensity of stimulation and this interaction is more pronounced for ascending trials since there were more errors for ascending trials at high intensities.

Some unpublished data involving magnitude estimations of the difference of pairs of visual stimuli also suggest the presence of this complex interaction. The stimuli and design of the experiment were identical to the Mullin, et al.

(1978) study described above. Although the geometric mean magnitude estimations of the difference between the ten pairs that were presented in both ascending and descending orders may be suspect because it cannot be assumed that the difference was always taken in the same direction by the subject, the means were 2.36 and 2.91 for ascending and descending trials respectively. More meaningful is the fact that the mean for trials in which the second stimulus was brighter was 2.50 while the mean for trials in which the second stimulus was dimmer was 2.99. This suggests that in the ascending condition, the second stimulus is seen as dimmer than, say, if viewed alone, and the subject judges the difference as smaller. In addition, the magnitude estimations of the difference between identical pairs of stimuli increased as absolute stimulus intensity increased. Once again, the evidence suggests that complex interactions underlie the data, interactions that are a function of the kind of trial and the absolute intensity of the stimuli.

Although the analysis of the frequency data has provided support for the bisymmetry axiom, the interaction between kind of trial and stimulus intensity noted for the two stimulus experiment appear to be a factor in the data from the present research involving the bisection paradigm using three stimuli. The difference between bisection points from ascending and descending trials suggests a violation of commutativity and/or reflexivity, although a

direct test of reflexivity was not possible. If it is correct that sequential and intensity effects are present in the data for a two stimulus task, it is reasonable to expect these factors to have an influence in a bisection task, an influence more complex than in a two stimulus situation. The next step is to investigate the quantitative nature of the relation between the obtained bisection points and the endpoint standards. To do so, however, requires certain assumptions. The assumptions necessary are that reflexivity is not violated and that the form of the psychophysical function is described by a power function. If these assumptions are correct, then $p + q = 1$ as in Equation 2, and Equation 9 can be fitted to the bisecction points estimated from the frequency data.

The obtained estimates of w from that expression were all greater than .50 indicating that the first endpoint presented was given more weight than the second. This suggests that the bisection point for descending trials should be greater than that for ascending trials, consistent with the linear least squares analysis of the frequency data. However, the function of Equation 9 did not fit the data well. A plot of the predicted against the estimated bisection points is given in Figure 10. The function tends to underestimate the bisection points for lower intensity stimuli. This implies that reflexivity is violated or that there are inaccuracies in the form of the psychophysical function that was assumed to describe the data. The estimates

of k, while similar to Stevens (1957) value of .33 for magnitude estimations of brightness, is larger than the value usually obtained from judgments of intervals (Curtis, 1970; Fagot & Stewart, 1970).

The Mullin, et al. (1978) study found that the power law for differences did not describe the data from reaction time to differences in pairs of brightnesses: there were interactions in the data involving the absolute intensities of the stimuli in a pair and the order in which these intensities were presented. Those differences in reaction time for ascending and descending trials are best interpreted as sequential order effects and not violations of reflexivity. The systematic deviations in Figure 10 would then seem to be attributable to these sequential effects as well as possible inaccuracies in the form of the psychophysical function assumed to describe the data.

The method of constant stimuli employed in the present experiment seems to avoid possible sources of bias that have been encountered by other researchers using the method of adjustment: biases due to the initial placement of the adjustable stimulus and other less definite sources (Garner, 1954; Fagot & Stewart, 1970). The bias due to ascending or descending order of presentation of the endpoints is still a factor, but that bias did not preclude the successful empirical verification of the bisymmetry axiom. Failure of many researchers to validate the bisection model through tests of non-parametric scalability should probably be

attributed to violations of commutativity, reflexivity or inaccuracies in the form of the psychophysical function assumed to describe the data (Fagot & Stewart, 1970) and not to violations of the bisymmetry axiom.

The reaction time data provide apparent convergent support for the bisymmetry axiom. The final bisection points, estimated from substituting paramater estimates obtained from fitting Equation 7 into Equation 10, were similar to each other and similar to the results from the frequency data. This finding is, however, open to question because of the poor fit of Equation 7. Perhaps this expression followed the monotone trend sufficiently well to locate its peak and, consequently, to closely estimate the bisection points, even though it did not predict accurately the actual observations.

It was initially proposed that the relation between reaction time and stimulus intensity is a reciprocal relation (Curtis, Paulos & Rule, 1973; Mullin & Curtis, 1973). In the present research, a reciprocal relation implies that reaction times should be an inverted U-shaped function of the variable stimuli for each pair of endpoint standards, and further, that the output exponent, m, of the psychophysical function should be negative. An inspection of the graphs of the reaction time data indicates that an inverse relation clearly is present. However, the estimate of the exponent m was positive in both of the fits of the psychophysical function assumed to underlie the data. In light of the

positive output exponent, the negative coefficient and the large additive constant, it seems apparent that the form of the psychophysical function assumed here is inadequate. Mullin, Curtis and Rule (1978) also found that while a reciprocal relation held between reaction time and stimulus intensity, there were complex interactions present in the data that could not be described by a psychophysical function similar in rationale and origin as the one assumed here. If interactions are present in the Mullin, et al. two stimulus task, it seems reasonable to expect interactions in the current three stimulus task, where the subject must judge the difference between pairs of intervals rather than the difference between pairs of stimuli.

It is also worth noting that the reaction time data seems to suggest violations of commutativity in a manner similar to the frequency data. An inspection of the graphs reveals that the curves generally peak at different points for ascending and descending trials, and the peak for descending trials is at a higher stimulus intensity than for ascending trials. This converges with the findings from the frequency data that descending bisection points are higher than ascending bisection points.

REFERENCES

Beck, J., & Shaw W. A. Discrimination of loudness similarity. Perception & Psychophysics, 1968, 3(2A), 105-107.

Coombs, C., Dawes, R. M., & Tversky, A. Mathematical Psychology, Prentice-Hall:Englewood Cliffs.

Cross, D. W. An application of mean value theory to psychological measurement. Unpublished manuscript, University of Michigan, 1965.

Curtis, D. W. Magnitude estimations and category judgments of brightness and brightness intervals: A two-stage interpretation. Journal of Experimental Psychology, 1970, 83(2), 201-208.

Curtis, D. W., Attneave, F., & Harrington, T. L. A test of a two stage model of magnitude judgment. Perception & Psychophysics, 1968, 3(1a), 25-31.

Curtis, D. W., Paulos, M. A., & Rule, S. J. Relation between disjunctive reaction time and stimulus difference. Journal of Experimental Psychology, 1973, 99(2), 167-173.

Fagot, R. F., & Stewart, M. R. Test of a response bias model of bisection. Perception & Psychophysics, 1970, 7(5), 257-262.

Gage, F. H. An experimental investigation of the measurability of auditory sensation. Proceedings of the Royal Society (London), 1934, 116B, 103-122, (a).

Gage, F. H. An experimental investigation of the measurability

of visual sensation. Proceedings of the Royal Society (London), 1934, 116B, 123-128, (b).

Garner, W. R. A technique and a scale for loudness measurement. The Journal of the Acoustical Society of America, 1954, 26(1), 73-88.

Hartley, H. E. The modified Gause-Newton method for the fitting of non-linear regression functions by least squares. Technometrics, 1961, 3, 269-280.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. Foundations of Measurement, Volume I. Academic Press: New York, 1971.

Marks, L. E. Sensory Processes: The New Psychophysics, Academic Press:New York, 1974.

Mc Gill, W. J. Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of Mathematical Psychology, Vol. 1, New York:Wiley, 1963.

Mullin, L. C., Curtis, D. W., & Rule, S. J. Disjunctive reaction time to simultaneously and sequentially presented visual stimuli. Paper presented at Western Psychological Association Convention, San Francisco, California, 1978.

Mullin, L. C., & Curtis, D. W. Simple and disjunctive reaction time as a function of loudness. Paper presented at Western Psychological Association Convention, Anaheim, California, 1973.

Nissen, M. J., & Pokorny, J. Wavelength effects on simple reaction time. Perception & Psychophysics, 1977, 22(5), 457-462.

Newman, E. B., Volkmann, J., & Stevens, S. S. On the method of bisection and its relation to a loudness scale. Americal Journal of Psychology, 1937, 49, 134-137.

Pfanzagl, J. Theory of Measurement. Wiley & Sons:New York, 1968.

Rule, S. J., Curtis, D. W., & Markley, R. P. Input and output transformations from magnitude estimation. Journal of Experimental Psychology, 1970, 86(3), 343-349.

Stevens, S. S. On the psychophysical law. Psychological Review, 1957, 64(3), 153-181.

Stevens, S. S. Decibels of light and sound. Physics Today, 1955, 8, 12-17.

Stewart, M. R., Fagot, R. F., & Eskildsen, P. R. Invariance tests for bisection and fractionation scaling. Perception & Psychophysics, 1967, 2(8), 323-327.

Weiss, D. W. Quantifying private events: A functional measurement analysis of equisection. Perception & Psychophysics, 1975, 16(4), 351, 357.

Woodworth, R. S., & Schlosberg, H. Experimental Psychology, Revised, Holt, Rinehart & Winston:New York, 1954.

APPENDIX A

Instructions

This is an experiment on bisection. The bisection task is one in which you are given two stimulus intensities, one lower than the other, and you must find the stimulus that divides the interval into two equal parts. For example, if you are shown two endpoint stimuli, call them a and c, you must find another stimulus, call it b, which divides the interval (a,c) so that the interval (a,b) is equal to the interval (b,c).

In the present experiment you will not have to actually find the midpoint stimulus. Instead, three stimuli will be presented consisting of two endpoints and a stimulus near the midpoint. You must decide simply whether the third stimulus is above or below the midpoint.

The stimuli will be different brightnesses of a red circle of light. The first two stimuli that you will see will be the ones that you are to regard as the two endpoints. The third stimulus is the one about which you are to make a judgment; that is, whether it is above or below the midpoint of the two endpoints. The endpoints will be presented on each trial, but sometimes the brighter one will be presented first and sometimes the dimmer one will be presented first. The two endpoints that you will be working with for the first half of the session will be the same on each trial, but in the two different orders just

described. Then we will take a break and change to another set of endpoints.

Because your full attention must be directed at looking at the lights so that you will see all three of them, the circle of light will be set to a very dim level between trials and between the three stimuli, so that you will know where they will appear. Also, because your full attention is required, you will start the trial whenever you are ready by pushing the button on the side of the box in front of you.

After looking at the sequence of brightnesses, if you decide that the third stimulus is closer to the first endpoint I want you to press the button on the top of the box that is to the left. Notice I am refering to the endpoint that is presented first and not to the intensity of that stimulus. If you decide that the third stimulus is closer to the second endpoint I want you to press the button that is to the right. If you decide it is exactly at the midpoint, you must still select one of the buttons to push.

Since this is also a reaction time experiment, I want you to make a judgment and press the appropriate button as quickly as you can. There are no right or wrong responses, because I want your judgments to be based on your impression of whether the third stimulus is above or below the midpoint. I want you to try to work as quickly and accurately as you can, and try to take

about the same amount of time between trials.

If you do not see all three stimuli, or if you think one of the stimuli flickered, or if you want to change your response, or if you think anything unusual is occurring, please do not hesitate to inform me.

APPENDIX B

For subject KM, the midpoints from the pooled frequency data are given in Table 4. As was noted earlier, these results are anomalous: the final bisection points obtained are not in the interval represented by the endpoints. Both final bisection points are below the lowest endpoints.

Column (1) of Table 14 presents the midpoints obtained when the frequencies from ascending and descending trials were analyzed separately. Astericks indicate which midpoints are outside of the interval represented by the endpoints. The midpoints here are very erratic. Notice the large discrepancy between midpoints for ascending and descending trials for endpoint (b o d). Notice also that two of the obtained midpoints are negative, a result that defies perceptual interpretation.

The proportion of variance accounted for by the regression analysis is given in Column (2) of Table 14. Less than half of the proportions are above 0.8 with one as low as 0.08. Clearly linear regression cannot give a good account of the obtained frequencies.

A graph of the variable stimuli against the standard normal scores for the obtained frequencies, along with the obtained least squares regression line is presented in Figure 9. The format is the same as the other graphs for frequencies. A comparison of this graph with Figures 4 through 8 reveals

TABLE 14

Data for subject KM. Column (1) presents the midpoints in cd/m^2 obtained from the least squares analysis of the frequency data. Column (2) is the proportion of variance accounted for by the regression analysis. Column (3) presents the midpoints in Column (1) transformed into decibels. See Appendix B for a discussion of these data.

Endpoints		(1)	(2)	(3)
(a o b)	Ascending	1.54*	.688	61.88
	Descending	4.02	.770	66.04
	Pooled	2.72	.827	64.35
(c o d)	Ascending	41.29*	.466	76.16
	Descending	44.36*	.435	76.47
	Pooled	42.98*	.566	76.33
(a o c)	Ascending	4.11	.709	66.14
	Descending	6.64	.938	68.22
	Pooled	5.24	.851	75.17
(b o d)	Ascending	4.50*	.502	66.53
	Descending	32.91	.936	75.17
	Pooled	22.95	.944	73.61
(a o b) o (c o d)				
	Ascending	-7.75*	.604	
	Descending	5.65	.765	67.52
	Pooled	1.03*	.806	60.13
(a o c) o (b o d)				
	Ascending	-17.08*	.086	
	Descending	7.35	.708	68.66
	Pooled	3.01*	.831	64.79

the failure of a linear least squares analysis to adequately describe the data for this subject.

Column (3), Table 14 presents the obtained bisection points transformed into decibels in the same manner that was used for the other subjects. The differences between ascending and descending trials for the basic pairs of endpoints were 4.16, 0.31, 2.08 and 8.64 respectively. (Differences for the final bisections could not be obtained because it is not possible to transform negative numbers into decibels.) The mean of these differences is 3.79 which is much larger than the differences obtained for the other subjects.

Differences between final bisection points needed for a test of the bisymmetry axiom were -1.14 and -4.66 for descending and pooled trials. (The difference could not be obtained for ascending trials, becuase of the negative obtained midpoints.) These are much larger than the differences obtained for the other subjects.

The only consistency of this subject's data with the other subjects is that the ascending midpoints were all less than the descending midpints. Thus, despite the unusual nature of KM's data, this one aspect of the results supports the comments made in the discussion regarding the direction of the differences between ascending and descending trials.

The reaction time data for KM are graphed in Figure 16. These data are again noteworth for their anomalous nature. First of all, this subject's reaction times were considerably faster than for the other subjects by 100-200 milliseconds. Second, the plots are generally flat with only a few showing the expected concave downward trend that was expected. Some plots even suggest a concave upward trend which is consistent with the unusual results obtained from this subject.

What are the possible explanations for the data obtained from this subject? First of all, it should be pointed out that this subject was a highly experienced psychophysics subject whose data in the past had been quite satisfactory. It may have been that KM's tolerance for such tasks had been exceeded. Also, looking at the fast reaction times, with their general flat nature, along with the unusual nature of the obtained midpoints, it seems possible that KM was sacrificing accuracy ofdiscrimination for speed of response. While some subject achieved reaction times as short as KM's, they acheived them for those variable stimuli that were easier to discriminate: the very high or very low variable stimuli for a pair of endpoints. KM, on the other hand, very obten achieved longer reaction times for the very high or very low variable stimuli. An inspection of Figure 18 reveals long reaction times for variable stimuli that presumably should be easy to make a judgment about.

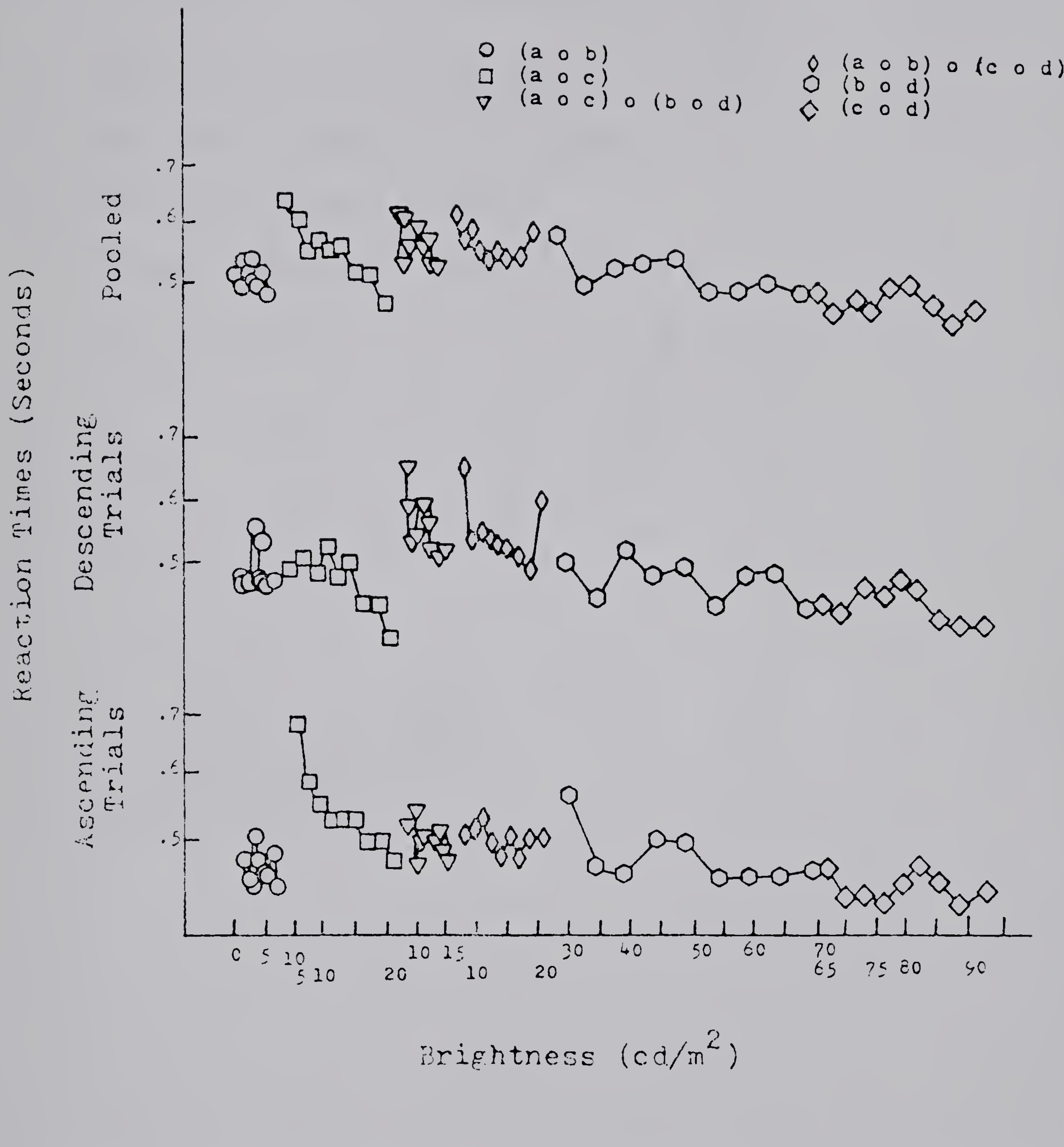


Figure 18: Reaction times as a function of brightness for subject KM. Lines connect reaction times for the variable stimuli associated with each pair of endpoints.

APPENDIX C

Equations

1. $F(a \circ b) = pF(a) + qF(b) + r$

2. $F(a \circ b) = pF(a) + (1 - p)F(b)$

3. $F(a \circ b) = \frac{1}{2}F(a) + \frac{1}{2}F(b)$

4. $RT_{abi} = f\left[d(a,i) - d(i,b)\right]^{m} + b$

5. $RT_{abi} = f\left[(Y_i - Y_a) - (Y_b - Y_i)\right]^{m} + b$

6. $RT_{abi} = a\left[(X_i^k - X_a^k) - (X_b^k - X_i^k)\right]^{m} + b$

7. $RT_{abi} = a\left[\left|X_a^k + X_b^k - 2X_i^k\right|\right]^{m} + b$

8. $z_i = aX_i + b$

9. $X_i = \left[wX_h^k + (1 - w)X_j^k\right]^{1/k}$

10. $X_i = \left[\frac{X_a^k + X_b^k}{2}\right]^{1/k}$